File S1

Mathematical derivations

Identity by descent (IBD) and identity by state (IBS)

Consider the process F_{g} in the main text. As we assume that the genes under consideration are neutral, the expected allele frequency for any individual equals to that in the ancestral generation, $E\left[x_{ijku}\right] = p_{ju}$. As x_{ijku} is an indicator function with values 0 or 1, it further holds that $E\left[x_{ijku}\, x_{ijku}\right] = p_{ju}$, and $E\left[x_{ijku'} x_{ijku}\right] = 0$ for $u \neq u'$. The probability that the alleles k and k' in a gene *j* are of the same type *u* for the individuals *i* and *i'* is given by $E\left[x_{ijku}\, x_{i'jk'u}\right]$. The probability that two randomly chosen alleles in a gene *j* are of the same type *u* for the individuals *i* and *i'* is given by $\sum_{k,k'} E\left[x_{ijku}\, x_{i'jk'u}\right]/4$. This probability can be decomposed into two components. First, the alleles may be identical by descent, the ancestral allele being of type *u*. The probability of this event is $\theta_{ii'} p_{ju}$. The second option is that the alleles are not identical by descent, but the alleles in the ancestral generation were both of type p_{ju}. As the probability of the latter event is $(1-\theta_{ii'})p_{ju}^2$, we obtain

$$\frac{1}{4}\sum_{k,k'} \mathrm{E}\left[x_{ijku}\, x_{i'jk'u}\right] = \theta_{ii'} p_{ju} + (1-\theta_{ii'})p_{ju}^2. \quad \text{(Eq. A1)}$$

Similarly, for $\neq u'$,

$$\frac{1}{4}\sum_{k,k'} \mathrm{E}\left[x_{ijku}\, x_{i'jk'u'}\right] = (1-\theta_{ii'})p_{ju}\, p_{ju'}. \quad \text{(Eq. A2)}$$

Thus,

$$\sum_{k,k'} \mathrm{Cov}[x_{ijku}\ , x_{i'jk'u'}\ ] = \left(\sum_{k,k'} \mathrm{E}\left[x_{ijku}\, x_{i'jk'u'}\right]\right) - 4p_{ju}\, p_{ju'} = 4\theta_{ii'} h_{j,u,u'}, \quad \text{(Eq. A3)}$$

where $h_{j,u,u'} = \delta_{u,u'} p_{ju} - p_{ju}\, p_{ju'}$. As we assume no linkage among the genes, it holds that $\mathrm{Cov}\left[x_{ijku}\, , x_{i'j'k'u'}\right] = 0$ for $j \neq j'$.

IBD and IBS based definitions for F_{ST}, F_{IS} and F_{IT}

We denote by f_0 the IBS probability for two alleles (same or different copy) in a randomly sampled gene *j* in a randomly sampled individual *i* in a randomly sampled subpopulation *X*. We denote by f_1 the IBS probability for

two alleles that are obtained by sampling two individuals (same or different) from a randomly chosen subpopulation, and then sampling alleles from a randomly chosen gene *j* of these two individuals. We denote by f_2 the IBS probability for two alleles that are obtained by randomly selecting two subpopulations (same or different), then two individuals from these, and then two alleles of the same randomly selected gene *j*.

We denote by $\omega = (1/n_{\mathcal{L}}) \sum_{j,u} p_{ju}^2$ the probability that two alleles from a randomly chosen gene *j* that are not identical by descent are identical by state. Note that IBS is possible when IBD does not hold although we ignore mutation, because multiple copies are available in the ancestral population. By Eq. A3 and the formulae for $\theta^{\mathcal{S}}$, θ, and $\theta^{\mathcal{P}}$ (see Table I in the main text),

$$f_0 = \frac{1}{n_{\mathcal{P}}} \sum_{X \in \mathcal{P}} \frac{1}{n_X} \sum_{i \in X} \frac{1}{n_{\mathcal{L}}} \sum_j \sum_u \frac{1}{4} \sum_{k,k'} \mathrm{E}\left[x_{ijku}\, x_{ij\,k'\,u}\right] = \theta^{\mathcal{S}} + \left(1 - \theta^{\mathcal{S}}\right)\, \omega, \quad \text{(Eq. A4)}$$

$$f_1 = \frac{1}{n_{\mathcal{P}}} \sum_{X \in \mathcal{P}} \frac{1}{n_X^2} \sum_{i,i' \in X} \frac{1}{n_{\mathcal{L}}} \sum_j \sum_u \frac{1}{4} \sum_{k,k'} \mathrm{E}\left[x_{ijku}\, x_{i'jk'u}\right] = \theta + (1-\theta)\omega, \quad \text{(Eq. A5)}$$

$$f_2 = \frac{1}{n_{\mathcal{P}}^2} \sum_{X,Y \in B} \frac{1}{n_X n_Y} \sum_{i \in X, i' \in Y} \frac{1}{n_{\mathcal{L}}} \sum_j \sum_u \frac{1}{4} \sum_{k,k'} \mathrm{E}\left[x_{ijku}\, x_{i'jk'u}\right] = \theta^{\mathcal{P}} + \left(1 - \theta^{\mathcal{P}}\right) \omega. \quad \text{(Eq. A6)}$$

We note that in the notation of COCKERHAM and WEIR (1987), $Q_1 = f_0$, $Q_2 = f_1$, $Q_3 = f_2$. COCKERHAM and WEIR (1987) defined $\sigma_0^2 = 1 - Q_1$ as the variance within individuals, $\sigma_1^2 = Q_1 - Q_2$ as the variance among individuals within subpopulations, and $\sigma_2^2 = Q_2 - Q_3$ as the variance among subpopulations within the population. Defining F_{ST} as the ratio of subpopulation variance to the total variance, we obtain

$$F_{ST} = \frac{\sigma_2^2}{\sigma_0^2 + \sigma_1^2 + \sigma_2^2} = \frac{f_1 - f_2}{1 - f_2} \quad \text{(Eq. A7)}$$

Similarly,

$$F_{IS} = \frac{f_0 - f_1}{1 - f_1}, \qquad F_{IT} = \frac{f_0 - f_2}{1 - f_2}.$$

measure the variance among individuals relative to that of the subpopulations or the total variance. Combining the above, we obtain the IBD-based analogues

$$F_{ST} = \frac{\theta + (1-\theta)\omega - \theta^{\mathcal{P}} - \left(1 - \theta^{\mathcal{P}}\right)\omega}{1 - \theta^{\mathcal{P}} - \left(1 - \theta^{\mathcal{P}}\right)\omega} = \frac{\left(\theta - \theta^{\mathcal{P}}\right)(1-\omega)}{(1-\theta^{\mathcal{P}})(1-\omega)} = \frac{\theta - \theta^{\mathcal{P}}}{1 - \theta^{\mathcal{P}}}, \quad \text{(Eq. A8)}$$

$$F_{IS} = \frac{\theta^{\mathcal{S}} - \theta}{1 - \theta}, \quad \text{(Eq. A9)}$$

$$F_{IT} = \frac{\theta^{\mathcal{S}} - \theta^{\mathcal{P}}}{1 - \theta^{\mathcal{P}}}. \quad \text{(Eq. A10)}$$

Additive genetic covariance among individuals

The covariance in the genetic value between traits *m* and m' in individuals *i* and i' is given by

$$\mathrm{Cov}[a_{im}, a_{i',m'}] = \sum_{j,k,u,k',u'} v_{jum}\, v_{ju'm'} \mathrm{Cov}\left[x_{ijku}\ , x_{i'jk'u'}\right] = 4\theta_{ii'} \sum_{j,u,u'} v_{jum}\, v_{ju'm'} h_{j,u,u'} \quad \text{(Eq. A11)}$$

To define the $\mathbf{G}^{\mathcal{A}}$ matrix, we construct a hypothetical individual representing the ancestral allele frequencies. We thus let x_{ju} be a set of random vectors (independently for each *j*) that follow the multinomial distribution with parameters 1 and p_{ju}, so that $\mathrm{E}\left[x_{ju}\right] = p_{ju}$. Let $\alpha = \left(\alpha_{1,} \ldots, \alpha_{n_{\mathcal{L}}}\right)$ be a random vector with $\alpha_m = \sum_{j,u} x_{ju}\, v_{jum}$. Let $\mathbf{G}^{\mathcal{A}}$ be the covariance matrix with $G^{\mathcal{A}}_{mm'} = 2\,\mathrm{Cov}[\alpha_m, \alpha_{m'}]$. We have

$$G^{\mathcal{A}}_{mm'} = 2 \sum_{j,u,u'} v_{jum} v_{ju'm'}\, \mathrm{Cov}\left[x_{ju}, x_{ju'}\right] = 2 \sum_{j,u,u'} v_{jum} v_{ju'm'} h_{j,u,u'} \quad \text{(Eq. A12)}$$

Combining the above shows that

$$\mathrm{Cov}[\boldsymbol{a}_i, \boldsymbol{a}_{i'}] = 2\theta_{ii'}\, \mathbf{G}^{\mathcal{A}} \quad \text{(Eq. A13)}$$

We may consider Eq. A14 as the definition for the $\mathbf{G}^{\mathcal{A}}$ matrix. $\mathbf{G}^{\mathcal{A}}$ depends on the allele frequencies in the ancestral population (p_{ju}) and additive values of the alleles (v_{ju}). Equivalently, as discussed above, we can define. $\mathbf{G}^{\mathcal{A}}$ by constructing a hypothetical individual *i* that represents the allele frequencies in the focal population, i.e. by randomizing the alleles according to the allele frequencies p_{ju} independently for each allele in each gene. This leads to $\theta_{ii} = 1/2$, so the interpretation that $\mathbf{G}^{\mathcal{A}}$ is the variance in the additive value in such a hypothetical individual is consistent with Eq. A13.

G in a local population

To compute $\mathbf{G}$ for a particular population, we apply Eq. A12 with the allele frequencies p_{ju} corresponding to those present in that population. We denote the amount of additive variance present in population X by $\mathbf{G}_X$, and the allele frequencies present in the population X by $p_{(X)ju} = \frac{1}{2n_X} \sum_{i \in X,k} x_{ijku}$.

By the above it holds that

$$\mathbf{G}_{X(m,m')} = 2\sum_{j,u,u'} v_{jum}\, v_{ju'm'}\left(\delta_{u,u'} p_{(X)ju} - p_{(X)ju}\, p_{(X)ju'}\right)$$
$$= 2\sum_{j}\left[\sum_{u} p_{(X)ju}\, v_{jum}\, v_{jum'} - \sum_{u,u'} p_{(X)ju}\, p_{(X)ju'}\, v_{jum}\, v_{ju'm'}\right]. \quad \text{(Eq. A14)}$$

The allele frequencies in the population X and thus also $\mathbf{G}_X$ depend on the realization of the process F_{g}. To compute the expectation of $\mathbf{G}_X$, we note that

$$\mathrm{E}[p_{(X)ju}] = \frac{1}{2n_X}\sum_{i\in X,k} \mathrm{E}[x_{ijku}] = p_{ju} \quad \text{(Eq. A15)}$$

and

$$\mathrm{E}[p_{(X)ju}\, p_{(X)ju'}] = \frac{1}{4n_X^2}\sum_{i,i'\in X,k,k'} \mathrm{E}[x_{ijku}\, x_{i'jk'u'}]. \quad \text{(Eq. A16)}$$

By Eqs. A2 and A3,

$$\mathrm{E}[p_{(X)ju}\, p_{(X)ju}] = \frac{1}{n_X^2}\sum_{i,i'\in X} \theta_{ii'} p_{ju} + (1-\theta_{ii'})p_{ju}^2 = \theta_X^{\mathcal{P}}\, p_{ju} + \left(1-\theta_X^{\mathcal{P}}\right) p_{ju}^2 \quad \text{(Eq. A17)}$$

and for $u \neq u'$

$$\mathrm{E}[p_{(X)ju}\, p_{(X)ju'}] = \frac{1}{n_X^2}\sum_{i,i'\in X} (1-\theta_{ii'}) p_{ju}\, p_{ju'} = (1-\theta_X^{\mathcal{P}})\, p_{ju}\, p_{ju'}. \quad \text{(Eq. A18)}$$

Thus

$$\sum_{u,u'} \mathrm{E}[p_{(X)ju}\; p_{(X)ju'}] v_{jum}\, v_{ju'm'} = \theta_X^{\mathcal{P}} \sum_{u} p_{ju}\, v_{jum}\, v_{jum'} + \left(1-\theta_X^{\mathcal{P}}\right)\sum_{u,u} p_{ju}\, p_{ju'}\, v_{jum}\, v_{ju'm'}, \quad \text{(Eq. A19)}$$

and the expectation of $\mathbf{G}_X$ is

$$\mathrm{E}[\mathbf{G}_{X(m,m')}] = \sum_{j}\left[\sum_{u} 2p_{ju}\, v_{jum}\, v_{jum'} - 2\theta_X^{\mathcal{P}} p_{ju}\, v_{jum}\, v_{jum'} - 2\left(1-\theta_X^{\mathcal{P}}\right)\sum_{u,u} p_{ju}\, p_{ju'}\, v_{jum}\, v_{ju'm'}\right]$$
$$= 2\left(1-\theta_X^{\mathcal{P}}\right)\sum_{j}\left[\sum_{u} p_{ju}\, v_{jum}\, v_{jum'} - \sum_{u,u} p_{ju}\, p_{ju'}\, v_{jum}\, v_{ju'm'}\right] = \left(1-\theta_X^{\mathcal{P}}\right) G_{mm'}^{\mathcal{A}}. \quad \text{(Eq. A20)}$$

We note in passing that $\mathbf{G}$ can be technically defined also for a "population" consisting of a single individual. Denoting $\mathbf{G}$ for individual *i* by $\mathbf{G}_i$, we have $\mathrm{E}[\mathbf{G}_i] = (1-\theta_{ii})\mathbf{G}^{\mathcal{A}}$. Thus, the expected amount of additive

variation present in a local population represented by a single non-inbred individual ($\theta_{ii} = 1/2$) is half of the additive variation that is present in the ancestral population.

Additive genetic variation among populations

The additive variance-covariance among the populations $\mathbf{D}$ is a random variable over the process F_{g}, and we next compute its expectation. We have

$$\mathrm{Cov}\left[\boldsymbol{a}_X^{\mathcal{P}}, \boldsymbol{a}_Y^{\mathcal{P}}\right] = \frac{1}{n_X n_X} \sum_{i\in X, j\in Y} \mathrm{Cov}[\boldsymbol{a}_i, \boldsymbol{a}_j] = \frac{1}{n_X n_Y} \sum_{i\in X, j\in Y} 2\theta_{ij}\, \mathbf{G}^{\mathcal{A}} = 2\theta_{XY}^{\mathcal{P}} \mathbf{G}^{\mathcal{A}}, \quad \text{(Eq. A21)}$$

$$\mathrm{Cov}\left[\boldsymbol{a}_X^{\mathcal{P}}, \boldsymbol{a}^{\mathcal{P}}\right] = \frac{1}{n_{\mathcal{P}}} \sum_{Y\in\mathcal{P}} \mathrm{Cov}\left[\boldsymbol{a}_X^{\mathcal{P}}, \boldsymbol{a}_Y^{\mathcal{P}}\right] = \frac{2}{n_{\mathcal{P}}} \sum_{Y\in\mathcal{P}} \theta_{XY}^{\mathcal{P}}\, \mathbf{G}^{\mathcal{A}}, \quad \text{(Eq. A22)}$$

$$\mathrm{Var}\left[\boldsymbol{a}^{\mathcal{P}}\right] = \frac{1}{n_{\mathcal{P}}} \sum_{X\in\mathcal{P}} \mathrm{Cov}[\boldsymbol{a}_X^{\mathcal{P}}, \boldsymbol{a}^{\mathcal{P}}] = \frac{2}{n_{\mathcal{P}}^2} \sum_{X,Y\in\mathcal{P}} \theta_{XY}^{\mathcal{P}}\, \mathbf{G}^{\mathcal{A}} = 2\theta^{\mathcal{P}} \mathbf{G}^{\mathcal{A}}. \quad \text{(Eq. A23)}$$

As the expectation of additive value is the same for each individual (and thus for $\boldsymbol{a}_X^{\mathcal{P}}$ and for $\boldsymbol{a}^{\mathcal{P}}$), it holds that

$$E[\mathbf{D}] = \frac{1}{n_{\mathcal{P}}} \sum_{X\in\mathcal{P}} \left(\mathrm{Var}\left[\boldsymbol{a}_X^{\mathcal{P}}\right] - 2\mathrm{Cov}[\boldsymbol{a}_X^{\mathcal{P}}, \boldsymbol{a}^{\mathcal{P}}] + \mathrm{Var}\left[\boldsymbol{a}^{\mathcal{P}}\right]\right) = \frac{1}{n_{\mathcal{P}}} \sum_{X\in\mathcal{P}} \left(2\theta_X^{\mathcal{P}} - \frac{4}{n_{\mathcal{P}}} \sum_{Y\in\mathcal{P}} \theta_{XY}^{\mathcal{P}} + 2\theta^{\mathcal{P}}\right) \mathbf{G}^{\mathcal{A}}$$
$$= 2\left(\theta - \theta^{\mathcal{P}}\right) \mathbf{G}^{\mathcal{A}}. \quad \text{(Eq. A24)}$$

Covariance structure of breeding experiment data

The four random processes ($F_{\mathrm{g}}, F_{\mathrm{g}'}, F_{\mathrm{s}}, F_{\mathrm{e}}$) described in the main text are disjoint, so that each of them can be applied in isolation or in combination with the others. We denote expectations (and related variances and covariances) over combinations of these random processes by listing the processes over which the expectations are taking in the subscript, e.g. $\mathrm{Cov}_{(\mathrm{s,e})}[\cdot]$. When an expectation is taken over some of the random processes while others are kept at a fixed realization, the result is conditional on the realizations kept fixed.

We assume that the environmental effects are independent between individuals, and are independent of additive genetic effects. This gives $\mathrm{Cov}_{(\mathrm{e})}\left[\boldsymbol{e}_i, \boldsymbol{e}_j\right] = \delta_{ij}\, \mathbf{V}_{\mathrm{E}}$, where $\mathbf{V}_{\mathrm{E}}$ denotes the environmental covariance matrix. The definition in left-hand side is conditional on the realizations of the processes $F_{\mathrm{g}}, F_{\mathrm{g}'}, F_{\mathrm{s}}$, but in this case the result in the right-hand side of is independent of these realizations. Similarly, we have $\mathrm{Cov}_{(\mathrm{g})}\left[\boldsymbol{a}_i, \boldsymbol{e}_j\right] = \mathrm{Cov}_{(\mathrm{g}')}\left[\boldsymbol{a}_i, \boldsymbol{e}_j\right] = \mathrm{Cov}_{(\mathrm{s})}\left[\boldsymbol{a}_i, \boldsymbol{e}_j\right] = 0$.

To proceed, we need to introduce some more notation. We denote co-ancestry coefficients for individual-population pairs as $\theta_{iY}^{i\mathcal{P}} = \frac{1}{n_Y}\sum_{j\in Y}\theta_{ij}$, so that $\theta_{XY}^{\mathcal{P}} = \frac{1}{n_X}\sum_{i\in X}\theta_{iY}^{i\mathcal{P}}$. We denote by s($i$) and d($i$) the sire and the dam of the individual *i,* and recall that $S(i)$ and $D(i)$ denote the local populations from which the sire and the dam of the individual *i* originate.

As shown by Eq. A13, the covariance among genetic effects for a pair of individuals is given by $\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_i, \boldsymbol{a}_j] = 2\theta_{ij}\,\mathbf{G}^{\mathcal{A}}$, where *i* and *j* refer to any individuals in the field populations or the laboratory population. Thus, utilizing the recursive formula for coancestry coefficients, $2\theta_{ij} = \theta_{s(i)j} + \theta_{d(i)j}$, we obtain for laboratory individuals *i* and *j*

$$\begin{aligned}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_i, \boldsymbol{p}_j] &= \frac{1}{2n_{S(j)}}\sum_{j'\in S(j)}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_i, \boldsymbol{a}_{j'}] + \frac{1}{2n_{D(j)}}\sum_{j'\in D(j)}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_i, \boldsymbol{a}_{j'}]\\ &= \frac{1}{2n_{S(j)}}\sum_{j'\in S(j)} 2\theta_{ij'}\mathbf{G}^{\mathcal{A}} + \frac{1}{2n_{D(j)}}\sum_{j'\in D(j)} 2\theta_{ij'}\boldsymbol{G}^{\mathcal{A}}\\ &= \frac{1}{2n_{S(j)}}\sum_{j'\in S(j)}[\theta_{s(i)j'} + \theta_{d(i)j'}]\mathbf{G}^{\mathcal{A}} + \frac{1}{2n_{D(j)}}\sum_{j'\in D(j)}[\theta_{s(i)j'} + \theta_{d(i)j'}]\mathbf{G}^{\mathcal{A}}\\ &= \frac{\mathbf{G}^{\mathcal{A}}}{2}[\theta_{s(i)S(j)}^{i\mathcal{P}} + \theta_{d(i)S(j)}^{i\mathcal{P}} + \theta_{s(i)D(j)}^{i\mathcal{P}} + \theta_{d(i)D(j)}^{i\mathcal{P}}]. \qquad (\mathrm{Eq.\,A25})\end{aligned}$$

For the field individual *i'* and lab-individual *j,* it holds that

$$\begin{aligned}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_{i'}, \boldsymbol{p}_j] &= \frac{1}{2n_{S(j)}}\sum_{j'\in S(j)}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_{i'}, \boldsymbol{a}_{j'}] + \frac{1}{2n_{D(j)}}\sum_{j'\in D(j)}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_{i'}, \boldsymbol{a}_{j'}]\\ &= \frac{1}{n_{S(j)}}\sum_{j'\in S(j)}\theta_{i'j'}\mathbf{G}^{\mathcal{A}} + \frac{1}{n_{D(j)}}\sum_{j'\in D(j)}\theta_{i'j'}\mathbf{G}^{\mathcal{A}} = \left[\theta_{i'S(j)}^{i\mathcal{P}} + \theta_{i'D(j)}^{i\mathcal{P}}\right]\mathbf{G}^{\mathcal{A}}. \qquad (\mathrm{Eq.\,A26})\end{aligned}$$

Thus, we obtain for the laboratory individuals *i* and *j*

$$\begin{aligned}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{p}_i, \boldsymbol{p}_j] &= \frac{1}{2n_{S(i)}}\sum_{i'\in S(i)}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}\left[\boldsymbol{a}_{i'}, \boldsymbol{p}_j\right] + \frac{1}{2n_{D(i)}}\sum_{i'\in D(i)}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}\left[\boldsymbol{a}_{i'}, \boldsymbol{p}_j\right]\\ &= \frac{1}{2n_{S(i)}}\sum_{i'\in S(i)}\left[\theta_{i'S(j)}^{i\mathcal{P}} + \theta_{i'D(j)}^{i\mathcal{P}}\right]\mathbf{G}^{\mathcal{A}} + \frac{1}{2n_{D(i)}}\sum_{i'\in D(i)}\left[\theta_{i'S(j)}^{i\mathcal{P}} + \theta_{i'D(j)}^{i\mathcal{P}}\right]\mathbf{G}^{\mathcal{A}}\\ &= \frac{1}{2}[\theta_{S(i)S(j)}^{\mathcal{P}} + \theta_{S(i)D(j)}^{\mathcal{P}} + \theta_{D(i)S(j)}^{\mathcal{P}} + \theta_{D(i)D(j)}^{\mathcal{P}}]\mathbf{G}^{\mathcal{A}}. \qquad (\mathrm{Eq.\,A27})\end{aligned}$$

Combining the above, for lab-individuals *i* and *j* it further holds that

$$\begin{aligned}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{s}_i,\boldsymbol{p}_j] &= \mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_i,\boldsymbol{p}_j] - \mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{p}_i,\boldsymbol{p}_j] \\ &= \frac{1}{2}\left[\theta^{i\mathcal{P}}_{s(i)S(j)} + \theta^{i\mathcal{P}}_{d(i)S(j)} + \theta^{i\mathcal{P}}_{s(i)D(j)} + \theta^{i\mathcal{P}}_{d(i)D(j)}\right]\mathbf{G}^{\mathcal{A}} \\ &\quad -\frac{1}{2}\left[\theta^{\mathcal{P}}_{S(i)S(j)} + \theta^{\mathcal{P}}_{S(i)D(j)} + \theta^{\mathcal{P}}_{D(i)S(j)} + \theta^{\mathcal{P}}_{D(i)D(j)}\right]\mathbf{G}^{\mathcal{A}} \quad \text{(Eq. A28)}\end{aligned}$$

and that

$$\begin{aligned}\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{s}_i,\boldsymbol{s}_j] &= \mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_i - \boldsymbol{p}_i, \boldsymbol{a}_j - \boldsymbol{p}_j] \\ &= \mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_i,\boldsymbol{a}_j] - \mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_i,\boldsymbol{p}_j] - \mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{p}_i,\boldsymbol{a}_j] + \mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{p}_i,\boldsymbol{p}_j] \\ &= \frac{1}{2}\theta_{ij}\,\mathbf{G}^{\mathcal{A}} - \frac{1}{2}\left[\theta^{i\mathcal{P}}_{s(i)S(j)} + \theta^{i\mathcal{P}}_{d(i)S(j)} + \theta^{i\mathcal{P}}_{s(i)D(j)} + \theta^{i\mathcal{P}}_{d(i)D(j)}\right]\mathbf{G}^{\mathcal{A}} \\ &\quad -\frac{1}{2}\left[\theta^{i\mathcal{P}}_{s(j)S(i)} + \theta^{i\mathcal{P}}_{d(j)S(i)} + \theta^{i\mathcal{P}}_{s(j)D(i)} + \theta^{i\mathcal{P}}_{d(j)D(i)}\right]\mathbf{G}^{\mathcal{A}} \\ &\quad +\frac{1}{2}\left[\theta^{\mathcal{P}}_{S(i)S(j)} + \theta^{\mathcal{P}}_{S(i)D(j)} + \theta^{\mathcal{P}}_{D(i)S(j)} + \theta^{\mathcal{P}}_{D(i)D(j)}\right]\mathbf{G}^{\mathcal{A}}. \quad \text{(Eq. A29)}\end{aligned}$$

We next treat the randomness associated to sampling individuals from the field populations to form the parental generation for the laboratory population. As we assume that the individuals used in the breeding design are a random sample from the field populations, we obtain for laboratory individuals *i* and *j*

$$\mathrm{E}_{(s)}\left[\theta^{i\mathcal{P}}_{s(i)S(j)}\right] = \mathrm{E}_{(s)}\left[\frac{1}{n_{S(j)}}\sum_{j'\in S(j)}\theta_{s(i)j'}\right] = \frac{1}{n_{S(i)}n_{S(j)}}\sum_{i'\in S(i),j'\in S(j)}\theta_{i'j'} = \theta^{\mathcal{P}}_{S(i)S(j)} \quad \text{(Eq. A30)}$$

Similarly,

$$\mathrm{E}_{(s)}\left[\theta^{i\mathcal{P}}_{s(i)D(j)}\right] = \theta^{\mathcal{P}}_{S(i)D(j)}, \qquad \mathrm{E}_{(s)}\left[\theta^{i\mathcal{P}}_{d(i)S(j)}\right] = \theta^{\mathcal{P}}_{D(i)S(j)}, \qquad \mathrm{E}_{(s)}\left[\theta^{i\mathcal{P}}_{d(i)D(j)}\right] = \theta^{\mathcal{P}}_{D(i)D(j)}. \quad \text{(Eq. A31)}$$

The expected sire-dam and dam-sire coancestry coefficients are given by

$$\mathrm{E}_{(s)}\left[\theta_{s(i)d(j)}\right] = \theta^{\mathcal{P}}_{S(i)D(j)}, \qquad \mathrm{E}_{(s)}\left[\theta_{d(i)s(j)}\right] = \theta^{\mathcal{P}}_{D(i)S(j)}. \quad \text{(Eq. A32)}$$

In case of sire-sire (or dam-dam) combinations, the two sires (dams) may represent the same individual or different individuals. Whether they represent the same or different individuals is part of the study design and thus independent of the sampling process, giving

$$\mathrm{E}_{(s)}\left[\theta_{s(i)s(j)}\right] = \begin{cases}\theta^{\mathcal{P}}_{S(i)S(j)} & s(i) \neq s(j) \\ \theta^{\mathcal{S}}_{S(i)} & s(i) = s(j),\end{cases} \quad \text{(Eq. A33)}$$

and

$$\mathrm{E}_{(s)}\left[\theta_{d(i)d(j)}\right] = \begin{cases}\theta^{\mathcal{P}}_{D(i)D(j)} & d(i) \neq d(j) \\ \theta^{\mathcal{S}}_{D(i)} & d(i) = d(j).\end{cases} \quad \text{(Eq. A34)}$$

As

$$\theta_{ij} = \begin{cases} \frac{1}{2} + \frac{1}{2}\,\theta_{s(i)d(j)}, & i = j \\ \frac{1}{4}\left(\theta_{s(i)s(j)} + \theta_{s(i)d(j)} + \theta_{d(i)s(j)} + \theta_{d(i)d(j)}\right), & i \neq j, \end{cases} \quad \text{(Eq. A35)}$$

we obtain

$$\mathrm{E}_{(s)}[\theta_{ij}] = \begin{cases} \frac{1}{2} + \frac{1}{2}\,\theta^{\mathcal{P}}_{S(i)D(i)} & i = j \\ \frac{1}{4}\left[\theta^{\mathcal{P}}_{S(i)S(j)} + \theta^{\mathcal{P}}_{S(i)D(j)} + \theta^{\mathcal{P}}_{D(i)S(j)} + \theta^{\mathcal{P}}_{D(i)D(j)}\right] & i \neq j, s(i) \neq s(j), d(i) \neq d(j) \\ \frac{1}{4}\left[\theta^{\mathcal{S}}_{S(i)} + \theta^{\mathcal{P}}_{S(i)D(j)} + \theta^{\mathcal{P}}_{D(i)S(j)} + \theta^{\mathcal{P}}_{D(i)D(j)}\right] & i \neq j, s(i) = s(j), d(i) \neq d(j) \\ \frac{1}{4}\left[\theta^{\mathcal{P}}_{S(i)S(j)} + \theta^{\mathcal{P}}_{S(i)D(j)} + \theta^{\mathcal{P}}_{D(i)S(j)} + \theta^{\mathcal{S}}_{D(i)}\right] & i \neq j, s(i) \neq s(j), d(i) = d(j) \\ \frac{1}{4}\left[\theta^{\mathcal{S}}_{S(i)} + \theta^{\mathcal{P}}_{S(i)D(j)} + \theta^{\mathcal{P}}_{D(i)S(j)} + \theta^{\mathcal{S}}_{D(i)}\right] & i \neq j, s(i) = s(j), d(i) = d(j) \end{cases} \quad \text{(Eq. A36)}$$

We next are now ready to extend the covariances over $(\mathrm{g}, \mathrm{g}')$ to covariances over $(\mathrm{g}, \mathrm{g}', \mathrm{s})$. Since the sampling process F_{s} is independent of the processes F_{g} and $F_{\mathrm{g}'}$, it holds for any random variables A and B,

$$\begin{aligned} \mathrm{Cov}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[A, B] &= \mathrm{E}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[AB] - \mathrm{E}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[A]\mathrm{E}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[B] \\ &= \mathrm{E}_{(\mathrm{s})}\left[\mathrm{E}_{(\mathrm{g},\mathrm{g}')}[AB] - \mathrm{E}_{(\mathrm{g},\mathrm{g}')}[A]\mathrm{E}_{(\mathrm{g},\mathrm{g}')}[B]\right] + \mathrm{E}_{(\mathrm{s})}\left[\mathrm{E}_{(\mathrm{g},\mathrm{g}')}[A]\mathrm{E}_{(\mathrm{g},\mathrm{g}')}[B]\right] - \mathrm{E}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[A]\mathrm{E}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[B] \\ &= \mathrm{E}_{(\mathrm{s})}\left[\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}(A, B)\right] + \mathrm{E}_{(\mathrm{s})}\left[\mathrm{E}_{(\mathrm{g},\mathrm{g}')}[A]\mathrm{E}_{(\mathrm{g},\mathrm{g}')}[B]\right] - \mathrm{E}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[A]\mathrm{E}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[B] \quad \text{(Eq. A37)} \end{aligned}$$

For any realization of the process F_{s} and any laboratory individual *i*,

$$\mathrm{E}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{a}_i] = \mathrm{E}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{p}_i] = \mathrm{E}_{(\mathrm{g},\mathrm{g}')}[\boldsymbol{s}_i] = 0, \quad \text{(Eq. A38)}$$

and thus, for any laboratory individuals *i* and *j*,

$$\mathrm{Cov}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[\boldsymbol{p}_i, \boldsymbol{p}_j] = \mathrm{E}_{(\mathrm{s})}\left[\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}(\boldsymbol{p}_i, \boldsymbol{p}_j)\right] = \frac{1}{2}\left[\theta^{\mathcal{P}}_{S(i)S(j)} + \theta^{\mathcal{P}}_{S(i)D(j)} + \theta^{\mathcal{P}}_{D(i)S(j)} + \theta^{\mathcal{P}}_{D(i)D(j)}\right]\mathbf{G}^{\mathcal{A}}, \quad \text{(Eq. A39)}$$

$$\begin{aligned} \mathrm{Cov}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[\boldsymbol{a}_i, \boldsymbol{p}_j] &= \mathrm{E}_{(\mathrm{s})}\left[\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}(\boldsymbol{a}_i, \boldsymbol{p}_j)\right] = \mathrm{E}_{(\mathrm{s})}\left[\theta^{i\mathcal{P}}_{iS(j)} + \theta^{i\mathcal{P}}_{iD(j)}\right]\mathbf{G}^{\mathcal{A}} \\ &= \frac{1}{2}\left[\theta^{\mathcal{P}}_{S(i)S(j)} + \theta^{\mathcal{P}}_{S(i)D(j)} + \theta^{\mathcal{P}}_{D(i)S(j)} + \theta^{\mathcal{P}}_{D(i)D(j)}\right]\mathbf{G}^{\mathcal{A}} = \mathrm{Cov}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[\boldsymbol{p}_i, \boldsymbol{p}_j], \quad \text{(Eq. A40)} \end{aligned}$$

$$\mathrm{Cov}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[\boldsymbol{s}_i, \boldsymbol{p}_j] = \mathrm{Cov}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[\boldsymbol{a}_i - \boldsymbol{p}_i, \boldsymbol{p}_j] = 0, \quad \text{(Eq. A41)}$$

$$\mathrm{Cov}_{(\mathrm{g},\mathrm{g}',\mathrm{s})}[\boldsymbol{a}_i, \boldsymbol{a}_j] = \mathrm{E}_{(\mathrm{s})}\left[\mathrm{Cov}_{(\mathrm{g},\mathrm{g}')}(\boldsymbol{a}_i, \boldsymbol{a}_j)\right] = 2\mathrm{E}_{(\mathrm{s})}[\theta_{\mathrm{ij}}]\mathbf{G}^{\mathcal{A}}, \quad \text{(Eq. A42)}$$

and thus

$$\mathrm{Cov}_{(\mathrm{g},\mathrm{g}',s)}[\boldsymbol{s}_i,\boldsymbol{s}_j] = \mathrm{Cov}_{(\mathrm{g},\mathrm{g}',s)}[\boldsymbol{a}_i-\boldsymbol{p}_i,\boldsymbol{a}_j-\boldsymbol{p}_j] = \mathrm{Cov}_{(\mathrm{g},\mathrm{g}',s)}[\boldsymbol{a}_i,\boldsymbol{a}_j] - \mathrm{Cov}_{(\mathrm{g},\mathrm{g}',s)}[\boldsymbol{p}_i,\boldsymbol{p}_j]$$
$$= 2f_{ij}\,\mathbf{G}^{\mathcal{A}} \quad \text{(Eq. A43)}$$

where

$$f_{ij} = \begin{cases} \frac{1}{2} - \frac{1}{4}\left[\theta^{\mathcal{P}}_{S(i)S(i)} + \theta^{\mathcal{P}}_{D(i)D(i)}\right] & i = j \\ 0 & i \neq j, s(i) \neq s(j), d(i) \neq d(j) \\ \frac{1}{4}\left[\theta^{\mathcal{S}}_{S(i)} - \theta^{\mathcal{P}}_{S(i)S(i)}\right] & i \neq j, s(i) = s(j), d(i) \neq d(j) \\ \frac{1}{4}\left[\theta^{\mathcal{S}}_{D(i)} - \theta^{\mathcal{P}}_{D(i)D(i)}\right] & i \neq j, s(i) \neq s(j), d(i) = d(j) \\ \frac{1}{4}\left[\theta^{\mathcal{S}}_{S(i)} + \theta^{\mathcal{S}}_{D(i)} - \theta^{\mathcal{P}}_{S(i)S(i)} - \theta^{\mathcal{P}}_{D(i)D(i)}\right] & i \neq j, s(i) = s(j), d(i) = d(j) \end{cases} \quad \text{(Eq. A44)}$$

References

COCKERHAM, C. C., and B. S. WEIR, 1987 Correlations, descent measures - drift with migration and mutation. Proceedings of the National Academy of Sciences of the United States of America **84:** 8512-8514.

File S2

Parameter estimation

Given neutral molecular data and the breeding experiment data, we fitted a neutral model to the data using Bayesian inference, and then tested for deviation from neutrality using the approach described in the main paper. Here we describe the statistical model, the prior distributions and a Markov Chain Monte Carlo (MCMC) method that was used to parameterize the model.

The statistical model for neutral molecular markers

We model the neutral divergence of the $n_{\mathcal{P}}$ local populations from a common ancestral population as a mixture of $n_{\mathcal{L}}$ independent lineages. The model summarized in this section will be published later in more detail (KARHUNEN and OVASKAINEN 2011).

The allele frequencies in locus j in lineage k are distributed as

$$\boldsymbol{z}_{kj} \sim \text{Dirichlet}\left(a_k \boldsymbol{q}_j\right).$$

where $\boldsymbol{q}_j$ denotes the vector of allele frequencies in the ancestral population, and a_k measures the amount of drift experienced by lineage k. The allele frequencies in locus j in local population A are defined as a mixture of the lineage-specific frequencies,

$$\boldsymbol{p}_j^A = \sum_{k=1}^{n_L} \kappa_k^A \boldsymbol{z}_{kj}\,.$$

Here, we assume that the mixture weights κ_k^A sum up to unity over the lineages, $\sum_{k=1}^{n_L} \kappa_k^A = 1$, so that vector $\boldsymbol{p}_j^A$ is a proper frequency distribution. The genotype (in terms of the neutral marker loci) of each individual in a population A is a multinomial random variable of the allele frequencies, $x_{ij\cdot} \sim \text{Multinomial}\left(2, \boldsymbol{p}_j^A\right)$.

The population-population coancestry coefficients depend on the model parameters as

$$\theta_{AB}^{\mathcal{P}} = \sum_{k=1}^{n_L} \frac{\kappa_k^A \kappa_k^B}{a_k + 1}.$$

Note that $\theta_{AB}^{\mathcal{P}}$ is defined through probability of IBD for neutral loci, and thus, it does not depend on allele frequencies $\boldsymbol{q}_j$ in the ancestral generation.

The Directed Acyclic Graph (DAG) that describes the dependencies among model variables (including both neutral marker data and quantitative trait data) is shown in Supplementary Figure 1. While the framework develop in the paper and presented in the DAG allows for the inclusion of an arbitrary level of inbreeding (θ_X^S), we did not account for inbreeding in the present estimation scheme and thus assumed $\theta_X^S = 0.5(1 + \theta_X^P)$.

Prior distributions

To parameterize the model with Bayesian inference, prior distributions need to be defined for the primary parameters of the model: $\boldsymbol{\mu}$, $\mathbf{G}^{\mathcal{A}}$, $\mathbf{V}_{\mathrm{E}}$,$\boldsymbol{q}$, $\boldsymbol{a}$ and $\boldsymbol{\kappa}$. We assume the priors

$$\mu_m \sim \mathrm{N}(0,100),$$

$$\mathbf{G}^{\mathcal{A}} \sim \mathrm{Wishart}(\mathbf{I}_m\ (m+1)^{-1}, m+1),$$

$$\mathbf{V}_{\mathrm{E}} \sim \mathrm{Wishart}(\mathbf{I}_m\ (m+1)^{-1}, m+1),$$

$$\boldsymbol{q}_j \sim \mathrm{Dirichlet}\left(\boldsymbol{\beta}_j^q\right),$$

$$\log a_k \sim \mathrm{N}(\mu_a, \sigma_a^2),$$

$$\boldsymbol{\kappa}_A \sim \mathrm{Dirichlet}(\boldsymbol{\beta}_A^\kappa).$$

Here the indices j, k, and A refer to loci, lineages and subpopulations, respectively. In the numerical examples of this paper, we assume the values $\boldsymbol{\beta}_j^q = \mathbf{1}_{n_j}$, $\mu_a = 1$, $\sigma_a^2 = 2$. We set the number of lineages equal to the number of populations, and assume that lineage A makes the dominant contribution to population A, i.e. that the matrix $\boldsymbol{\kappa}$ is diagonally dominant. To do so, we let

$$\beta_{AA}^\kappa = 0.8, \text{ and } \beta_{Ak}^\kappa = \frac{0.2}{n_L - 1} \text{ for } k \neq A,$$

and truncate the prior by the requirement that $\beta_{AA}^\kappa > \beta_{Ak}^\kappa$ for all $k \neq A$. The latter specification ensures that label switching is not possible, which property improves the mixing of Markov Chain Monter Carlo (MCMC) algorithm (GELMAN *et al.* 2004).

The number of alleles n_j in neutral marker locus j in the ancestral generation is generally unknown, as some alleles may have disappeared after the lineages have diverged or are not present in the sampled individuals. Due to mathematical properties of the Dirichlet distribution, we may pool all such unobserved alleles into one 'meta-allele'. Thus, we define n_j as the number of distinct alleles observed in locus j plus one.

Parameter estimation through MCMC

Each model parameter was sampled from its full conditional while keeping the other parameters fixed. Most parameters were updated using a random-walk Metropolis-Hastings (MH) algorithm, the proposal distributions given below. For these the variances of the proposal distributions were adjusted during the burn-in following Ovaskainen et al. (2008) to give an accept ratio of 0.44 for a single parameter or 0.22 for a vector of parameters.

- Sampling the additive genetic variance-covariance matrix $\mathbf{G}^{\mathcal{A}}$ was conducted using the MH algorithm. We used the proposal distribution $\mathrm{Wishart}(\mathbf{G}^{\mathcal{A}}/\nu, \nu)$, where the parameter ν was adjusted during the burn-in.

- Sampling the residual variance-covariance matrix $\mathbf{V}_{\mathrm{E}}$ was conducted using the MH algorithm. We used the proposal distribution $\mathrm{Wishart}(\mathbf{V}_{\mathrm{E}}/\nu, \nu)$, where the parameter ν was adjusted during the burn-in.

- Sampling the overall mean $\boldsymbol{\mu}$ and the population specific means $\mathbf{A}^{\mathcal{P}}$. Conditional on the other parameters, the full conditional joint distribution of the parameters $(\boldsymbol{\mu}, \mathbf{A}^{\mathcal{P}})$ follows a multivariate normal distribution. These parameters were thus updated directly (not using the MH algorithm) by drawing a random deviate from the full conditional.

- Sampling drift parameters $\boldsymbol{\alpha}$. We used $\mathrm{N}(a_k, \delta_{a_k}^2)$ distributions separately for each k to draw proposals for $\log a_k$. The variance parameters $\delta_{a_k}^2$ were adjusted during the burn-in.

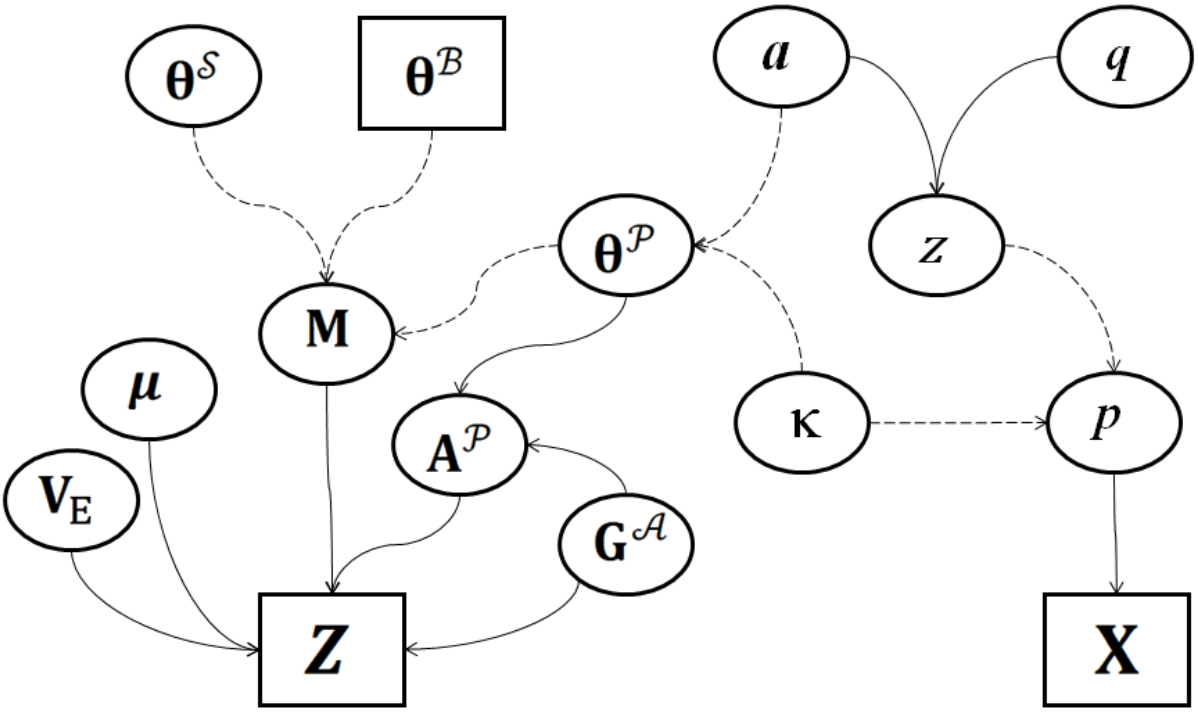
- Sampling lineage loadings $\boldsymbol{\kappa}$. We used truncated $\mathrm{Dirichlet}(\delta_{\boldsymbol{\kappa}_A} \boldsymbol{\kappa}_A)$ distributions separately for each A and j to draw proposals for $\boldsymbol{\kappa}_A$. The $\delta_{\boldsymbol{\kappa}_A}$'s are proposal parameters that were adjusted during the burn-in.

- Sampling ancestral allele frequencies $\boldsymbol{q}$. We used truncated $\mathrm{Dirichlet}(\delta_{q_j} \boldsymbol{q}_j)$ distributions separately for each j to draw proposals for $\boldsymbol{q}_j$. The δ_{q_j}'s are proposal parameters that were adjusted during the burn-in.

- Sampling allele frequencies $\boldsymbol{z}$. We used $\mathrm{Dirichlet}(\delta_{\boldsymbol{z}_{kj}} \boldsymbol{z}_{kj})$ distributions separately for each lineage k and locus j. The $\delta_{\boldsymbol{z}_{kj}}$'s parameters were adjusted during the burn-in.

We implemented the algorithm in Mathematica 7.1 (WOLFRAM RESEARCH 2008), and used 1,000 iterations for the burn-in and 2,000 for sampling the posterior distribution. We note that these sample sizes are rather small, and they were chosen due to logistic constrains (the estimation was conducted for 3,600 data sets in total). We tested for the sufficiency for mixing by examining the chains for convergence and for running a sample of the datasets for 10,000 + 20,000 iterations. Both tests (as well as the Results, which separated well the cases with and without selection, see main text) indicated that the length of the MCMC chain was sufficient for the present purpose. For the case of real data analysis, we naturally recommend a much longer MCMC chain.

Supplementary Figure 1. Directed acyclic graph (DAG) for the Bayesian model. The quantities in the ellipses are the parameter values being estimated. In the single boxes are the observations (the neutral data $\boldsymbol{X}$ and the quantitative trait data $\boldsymbol{Z}$) and in the double boxes are the fixed quantities. The full arrows give stochastic relationships (i.e. those that have probability distributions associated with them) and the dashed arrows are for deterministic relationships.



References

ALBERT, A. Y. K., S. SAWAYA, T. H. VINES, A. K. KNECHT, C. T. MILLER *et al.*, 2008 The genetics of adaptive shape shift in stickleback: Pleiotropy and effect size. Evolution **62:** 76-85.

GELMAN, A., J. B. CARLIN, H. S. STERN and D. B. RUBIN, 2004 *Bayesian Data Analysis*. Chapman & Hall/CRC, Boca Raton.

KARHUNEN, M., and O. OVASKAINEN, 2011 Defining and estimating FST through coancestry. Manuscript.

WOLFRAM RESEARCH, I., 2008 *Mathematica Edition: Version 7.1*. Wolfram Research, Inc., Champaign, Illinois.

File S3

Generation of individual-based data

We assumed that in each generation of each population the breeding population consists of 20 individuals. We assumed a hermaphrodite species and randomized the two parents for each propagule independently of each other (selfing allowed). To model gene flow among the populations, we assumed that the parent is an immigrant from one of the other populations with probability ε=0.01.

We assume a set of 32 loci determining the genotypic values among these traits, with 5 allelic variants in each locus, each having an equal frequency in the ancestral generation. We randomized the additive values of each allele in each locus from a two-dimensional multivariate normal distribution with mean zero and such a variance-covariance matrix (see Appendix A) that the expected amount of additive genetic variation in the ancestral generation was

$$\mathrm{E}[\mathbf{G}^{\mathcal{A}}] = \begin{pmatrix} 1.0 & 0.9 \\ 0.9 & 1.0 \end{pmatrix}.$$

The phenotypic value of individual i was modeled as the sum of the additive component and the environmental effect, i.e. $\mathbf{z}_i = \mathbf{a}_i + \mathbf{e}_i$, where the environmental effect $\mathbf{e}_i$ was randomized from the two dimensional multivariate normal distribution with mean zero and variance-covariance matrix

$$\mathbf{V}_E = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}.$$

If modeling traits under selection, we let the populations become adapted locally to their environments by assuming that the environment experienced by a population X favored an optimal phenotype $\boldsymbol{q}_X$. We assumed that the $\boldsymbol{q}_X$ values are distributed (i.i.d. among the populations) multinormally with mean zero and variance covariance matrix

$$\mathbf{V}_S = \begin{pmatrix} 5 & -4.5 \\ -4.5 & 5 \end{pmatrix},$$

and we thus have assumed that the expected direction of population divergence through the selective gradient is not proportional to $\mathbf{G}^{\mathcal{A}}$.

We assume that the populations undergo weak selection, so that the initial number of propagules produced is high (ten times the actual population size), and selection operates through competition influencing survival. We define the fitness of an individual with phenotype $\mathbf{z}$ in population j as

$$W_j(\mathbf{z}) = \exp(-(\mathbf{z} - \boldsymbol{q}_j)\mathbf{V}_F{}^{-1}(\mathbf{z} - \boldsymbol{q}_j))$$

where the variance-covariance matrix

$$\mathbf{V}_F = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$$

describes the shape of the fitness landscape around the local optimum. We selected the 20 individuals that survived to form next generation among 200 propagules by sampling each individual with a probability that was proportional to its fitness value. In the case of no selection, the weights were set equal for all individuals.

To generate phenotypic data that allows for the estimation of genetic variability among and between populations, we assumed a breeding design in which five pairs of individuals were randomly sampled from each of the eight populations. We assumed a full-sib design with five offspring for each of the pairs, so that the total number of offspring was 8 x 5 x 5 = 200. These individuals were assigned environmental effects as for the wild individuals, and their phenotypes were then recorded.

For neutral molecular markers, we also assumed a set of 32 loci with 5 alleles in each locus with equal frequencies in the ancestral population. We randomized the flow of the neutral alleles from the ancestral to the present generation, sampled ten individuals from each population, and genotyped these for the neutral molecular markers.

File S4

Implementation of the method of Martin et al. (2008)

We analyzed the simulated data sets with Martin et al.'s method (MARTIN *et al.* 2008b). This method involves two separate tests: Comparing the estimates of the intra-population G matrix and the covariance matrix of population means for proportionality, and secondly, testing the proportionality coefficient of these matrices against the F_{ST} of molecular markers. The first test should see selection that has a magnitude comparable to random drift, but a direction incompatible with $\mathbf{G}^{\mathcal{A}}$. The second test should see selection that has a magnitude incompatible with random drift, whether balancing or disruptive. Both tests are based on the assumption that

$$\mathrm{E}[\mathbf{D}] = \frac{2F_{ST}}{1 - F_{ST}} \mathrm{E}[\mathbf{G}_X]$$

under neutrality. We performed these tests following the procedure described by the authors (CHAPUIS *et al.* 2008; MARTIN *et al.* 2008b). Thus, we estimated the G matrices using R's MANOVA with three covariance components: inter-population, inter-family and inter-individual. The inter-population covariance matrix was used as an estimate of $\mathrm{E}[\mathbf{D}]$, whereas the inter-population covariance matrix was multiplied by two to yield an estimate of $\mathrm{E}[\mathbf{G}_X]$, which is appropriate for full-sib study design (LYNCH and WALSH 1998). Because the phenotypic data was known to be multivariate normal, it was not Cox transformed, but we followed the recommendation of standardization (CHAPUIS *et al.* 2008).

After this, the proportionality of $\mathrm{E}[\mathbf{D}]$ and $\mathrm{E}[\mathbf{G}_X]$ was tested, and a 95 % confidence interval was derived for the proportionality coefficient using code provided by MARTIN *et al.* (2008a). The Bartlett *p* value of the proportionality test was recorded. To perform the second test of MARTIN *et al.* (2008b), a 95 % confidence interval was estimated for the F_{ST} using bootstrap over loci and the Weir-Cockerham estimator (WEIR and COCKERHAM 1984) implemented in Fstat (GOUDET 1995). The confidence interval of F_{ST} was transformed into the confidence interval of $\frac{2F_{ST}}{1-F_{ST}}$. This was compared with the confidence interval of the proportionality coefficient obtained from the phenotypic data. If these intervals did not overlap, this was taken as a signal of selection.

References

CHAPUIS, E., G. MARTIN and J. GOUDET, 2008 Effects of Selection and Drift on G Matrix Evolution in a Heterogeneous Environment: A Multivariate Q(st)-F-st Test With the Freshwater Snail Galba truncatula. Genetics **180:** 2151-2161.

GOUDET, J., 1995 FSTAT (Version 1.2): A computer program to calculate F-statistics. Journal of Heredity **86:** 485-486.

LYNCH, M., and B. WALSH, 1998 *Genetics and analysis of quantitative traits*. Sinauer Associates Incorporated, New York.

MARTIN, G., E. CHAPUIS and J. GOUDET, 2008a http://www.isem.cnrs.fr/spip.php?article934, pp. Institut des sciences de l'evolution montpellier, Montpellier.

MARTIN, G., E. CHAPUIS and J. GOUDET, 2008b Multivariate Q(st)-F-st Comparisons: A Neutrality Test for the Evolution of the G Matrix in Structured Populations. Genetics **180:** 2135-2149.

WEIR, B. S., and C. C. COCKERHAM, 1984 Estimating F-Statistics for the Analysis of Population-Structure. Evolution **38:** 1358-1370.