

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, $E[x_{ijk_u}] = p_{ju}$. As x_{ijk_u} is an indicator function with values 0 or 1, it further holds that $E[x_{ijk_u} x_{ijk_u}] = p_{ju}$, and $E[x_{ijk_u} x_{ijk_u}] = 0$ 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 $E[x_{ijk_u} x_{i'jk'_u}]$. 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 $\sum_{k,k'} E[x_{ijk_u} x_{i'jk'_u}] / 4$. 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

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

Similarly, for $\neq u'$,

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

Thus,

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

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 $\text{Cov}[x_{ijk_u}, x_{i'j'k'_u'}] = 0$ 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 θ^S , θ , and θ^P (see Table I in the main text),

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

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

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

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,

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

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

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

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

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

Additive genetic covariance among individuals

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

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

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 $E[x_{ju}] = p_{ju}$. Let $\alpha = (\alpha_1, \dots, \alpha_{n_L})$ be a random vector with $\alpha_m = \sum_{j,u} x_{ju} v_{jum}$. Let $\mathbf{G}^{\mathcal{A}}$ be the covariance matrix with $G_{mm'}^{\mathcal{A}} = 2 \text{Cov}[\alpha_m, \alpha_{m'}]$. We have

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

Combining the above shows that

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

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 $p_{(X)ju} = \frac{1}{2n_X} \sum_{i \in X, k} x_{ijk u}$.

By the above it holds that

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

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

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

and

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

By Eqs. A2 and A3,

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

and for $u \neq u'$

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

Thus

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

and the expectation of \mathbf{G}_X is

$$\begin{aligned} E[\mathbf{G}_{X(m,m')}] &= \sum_j \left[\sum_u 2p_{ju} v_{jum} v_{ju'm'} - 2\theta_X^P p_{ju} v_{jum} v_{jum'} - 2(1 - \theta_X^P) \sum_{u,u'} p_{ju} p_{ju'} v_{jum} v_{ju'm'} \right] \\ &= 2(1 - \theta_X^P) \sum_j \left[\sum_u p_{ju} v_{jum} v_{ju'm'} - \sum_{u,u'} p_{ju} p_{ju'} v_{jum} v_{ju'm'} \right] = (1 - \theta_X^P) \mathbf{G}_{mm'}^A. \quad (\text{Eq. A20}) \end{aligned}$$

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 $E[\mathbf{G}_i] = (1 - \theta_{ii}) \mathbf{G}^A$. 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

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

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

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

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

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

Covariance structure of breeding experiment data

The four random processes ($F_g, F_{g'}, F_s, F_e$) 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. $\text{Cov}_{(s,e)}[\cdot]$. 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 $\text{Cov}_{(e)}[\mathbf{e}_i, \mathbf{e}_j] = \delta_{ij} \mathbf{V}_E$, where \mathbf{V}_E denotes the environmental covariance matrix. The definition in left-hand side is conditional on the realizations of the processes $F_g, F_{g'}, F_s$, but in this case the result in the right-hand side of is independent of these realizations. Similarly, we have $\text{Cov}_{(g)}[\mathbf{a}_i, \mathbf{e}_j] = \text{Cov}_{(g')}[\mathbf{a}_i, \mathbf{e}_j] = \text{Cov}_{(s)}[\mathbf{a}_i, \mathbf{e}_j] = 0$.

To proceed, we need to introduce some more notation. We denote co-ancestry coefficients for individual-population pairs as $\theta_{iY}^{iP} = \frac{1}{n_Y} \sum_{j \in Y} \theta_{ij}$, so that $\theta_{XY}^P = \frac{1}{n_X} \sum_{i \in X} \theta_{iY}^{iP}$. 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 $\text{Cov}_{(g,g')}[\mathbf{a}_i, \mathbf{a}_j] = 2\theta_{ij} \mathbf{G}^{\mathcal{A}}$, 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

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

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

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

Thus, we obtain for the laboratory individuals i and j

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

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

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

and that

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

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

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

Similarly,

$$\mathbb{E}_{(s)}[\theta_{s(i)D(j)}^{iP}] = \theta_{S(i)D(j)}^P, \quad \mathbb{E}_{(s)}[\theta_{d(i)S(j)}^{iP}] = \theta_{D(i)S(j)}^P, \quad \mathbb{E}_{(s)}[\theta_{d(i)D(j)}^{iP}] = \theta_{D(i)D(j)}^P. \quad (\text{Eq. A31})$$

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

$$\mathbb{E}_{(s)}[\theta_{s(i)d(j)}] = \theta_{S(i)D(j)}^P, \quad \mathbb{E}_{(s)}[\theta_{d(i)s(j)}] = \theta_{D(i)S(j)}^P. \quad (\text{Eq. A32})$$

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

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

and

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

As

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

we obtain

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

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

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

For any realization of the process F_s and any laboratory individual i ,

$$E_{(g, g')}[\mathbf{a}_i] = E_{(g, g')}[\mathbf{p}_i] = E_{(g, g')}[\mathbf{s}_i] = 0, \quad (\text{Eq. A38})$$

and thus, for any laboratory individuals i and j ,

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

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

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

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

and thus

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

where

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

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.