Theor Appl Genet (1991) 81:233–238

On the principle underlying the tabular method to compute coancestry *

H. L. Chang, R. L. Fernando ** and M. Grossman

Department of Animal Sciences, 1207 W. Gregory Drive, University of Illinois, Urbana, IL 61801, USA

Received December 20, 1989; Accepted June 12, 1990
Communicated by E. J. Eisen

Summary. The tabular method to compute coancestry between two individuals is based on the principle that coancestry may be computed as the average coancestry between one individual and the parents of the other, on the condition that the former individual is not a direct descendent of the latter. It follows that coancestry also may be computed as the average of the four coancestries between the parents of the two individuals, on the condition that each individual is not a direct descendent of the other. The requirement for these conditions is explained.

Key words: Coancestry – Tabular method

Introduction

The tabular method was developed as a systematic procedure to compute coefficient of inbreeding (Emik and Terrill 1949). Now, it is used widely by breeders to construct the matrix of additive relationships (twice coancestry) between relatives for autosomal loci (Pirchner 1983; Van Vleck et al. 1987). A tabular method to construct the matrix of covariances between relatives for X-chromosomal loci also has been developed (Fernando and Grossman 1990). The inverse of these matrices can be used in the mixed model equations to predict breeding values by best linear unbiased prediction (Henderson 1976; Fernando and Grossman 1990).

The tabular method to construct the matrix of coancestries for autosomal loci is based on the principle that coancestry between two individuals may be expressed as the average coancestry between one individual and the parents of the other. This principle is always true, with the condition that *the former individual is not a direct descendent of the latter* (Emik and Terrill 1949). In the tabular method, this condition is always met because individuals are ordered such that progeny follow parents (Van Vleck et al. 1987). However, this principle for computing coancestry is often given without explicitly stating this condition (e.g., see p. 75, Kempthorne 1969; p. 70, Pirchner 1983; p. 123, Doolittle 1987; p. 205, Van Vleck et al. 1987; p. 81, Falconer 1989), perhaps because a clear explanation for this condition is not widely known. The object of this paper is to provide such an explanation.

Theory

Let X and Y be two individuals and let M be the maternal parent and P the paternal parent of Y (Fig. 1). Individuals and their genotypes at a given autosomal locus are summarized below:

Individual	M	P	X	Y
Genotype	$m_m\, m_p$	$p_m\, p_p$	$x_m\, x_p$	$y_m\, y_p$
Random gene	m	p	x	y

where the subscript m denotes a gene of maternal origin and the subscript p denotes a gene of paternal origin.

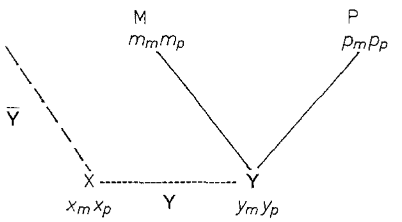

Fig. 1. Paths of inheritance for M, P, X, and Y

---

* This research was supported in part by the Illinois Agricultural Experiment Station, Hatch Projects 35-345 and 35-367

** To whom offprint requests should be addressed

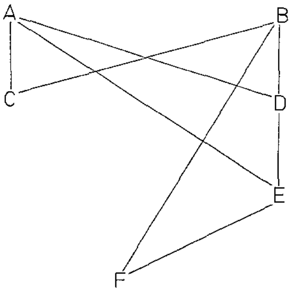

Fig. 2. Pedigree (Fig. 5.1 from Kempthorne 1969)

Coancestry between two arbitrary individuals X and Y (r_{XY}) is the probability that a random gene from the locus in X(x) is identical by descent ($\equiv$) to a random gene from that locus in Y(y) (Malécot 1969). Now, x can be identical by descent to y in two mutually exclusive ways:

> y could be the maternal gene in Y ($y_m = y$), with probability $\frac{1}{2}$, and y_m is identical by descent to x

or

> y could be the paternal gene in Y ($y_p = y$), with probability $\frac{1}{2}$, and y_p is identical by descent to x,

so that

$$\begin{aligned} r_{XY} = \mathrm{P}(x \equiv y) &= \mathrm{P}(x \equiv y_m, y_m = y) + \mathrm{P}(x \equiv y_p, y_p = y) \\ &= \mathrm{P}(x \equiv y_m | y_m = y)\ \mathrm{P}(y_m = y) \\ &\quad + \mathrm{P}(x \equiv y_p | y_p = y)\ \mathrm{P}(y_p = y) \\ &= \tfrac{1}{2}[\mathrm{P}(x \equiv y_m | y_m = y) + \mathrm{P}(x \equiv y_p | y_p = y)] \\ &= \tfrac{1}{2}[\mathrm{P}(x \equiv y_m) + \mathrm{P}(x \equiv y_p)] \end{aligned} \tag{1}$$

because the process of sampling the random gene y from Y is independent of genes in X being identical by descent to genes in Y. Equation 1 is true, irrespective of whether one individual is a direct descendent of the other.

Because y_m and m are random genes from M, and because y_p and p are random genes from P, it is tempting to write Eq. 1 as

$$r_{XY} = \mathrm{P}(x \equiv y) = \tfrac{1}{2}[\mathrm{P}(x \equiv m) + \mathrm{P}(x \equiv p)] = \tfrac{1}{2}[r_{XM} + r_{XP}] \tag{2}$$

where r_{XM} is the coancestry between X and M and r_{XP} is the coancestry between X and P. If X is not a direct descendent of Y, Eq. 2 is always true. If X is a direct decendent of Y, however, Eq. 2 is not always true because $\mathrm{P}(x \equiv y_m)$ is greater than or equal to $\mathrm{P}(x \equiv m)$ (see Eq. 4a) and $\mathrm{P}(x \equiv y_p)$ is greater than or equal to $\mathrm{P}(x \equiv p)$ (see Eq. 4b).

The random gene x can be identical by descent to y through two mutually exclusive sets of paths of inheritance: (1) through the set of paths in which X inherits genes from Y (denoted Y; or (2) through the remaining set of paths (denoted $\bar{\mathrm{Y}}$). For example, let Y = E and X = F (Fig. 2). Then the path directly from E to F is Y and the path from E to F through the common ancestor B is $\bar{\mathrm{Y}}$.

Thus, coancestry between any X and Y can be written as

$$\mathrm{P}(x \equiv y) = \mathrm{P}(x \equiv y, \mathrm{Y}) + \mathrm{P}(x \equiv y, \bar{\mathrm{Y}}) \tag{3}$$

or

$$r_{XY} = r_{XY,Y} + r_{XY,\bar{Y}}$$

Similarly, the two probabilities in Eq. 1 can be written as

$$\mathrm{P}(x \equiv y_m) = \mathrm{P}(x \equiv y_m, \mathrm{Y}) + \mathrm{P}(x \equiv y_m, \bar{\mathrm{Y}}) \tag{3a}$$

and

$$\mathrm{P}(x \equiv y_p) = \mathrm{P}(x \equiv y_p, \mathrm{Y}) + \mathrm{P}(x \equiv y_p, \bar{\mathrm{Y}})\ . \tag{3b}$$

For the first term in Eq. 3a, Y contains the path from M to X through Y. Therefore, the random gene y_m is one generation closer than m to X, so that $\mathrm{P}(x \equiv y_m, \mathrm{Y})$ is greater than or equal to $\mathrm{P}(x \equiv m, \mathrm{Y})$, even though m and y_m are random genes from M. Proof that

$$\begin{aligned} &\mathrm{P}(x \equiv y_m, \mathrm{Y}) = 2r_{XM,Y} - \mathrm{P}(x = y, \mathrm{Y})\,[F_M + F_Y] \\ &\geq \mathrm{P}(x \equiv m, \mathrm{Y}) = r_{XM,Y} \end{aligned} \tag{4a}$$

is in Appendix A, which should be read after Appendix B for the better understanding.

Similarly, the first term in Eq. 3b is

$$\begin{aligned} &\mathrm{P}(x \equiv y_p, \mathrm{Y}) = 2r_{XP,Y} - \mathrm{P}(x = y, \mathrm{Y})\,[F_P + F_Y] \\ &\geq \mathrm{P}(x \equiv p, \mathrm{Y}) = r_{XP,Y} \end{aligned} \tag{4b}$$

Now, the same reasoning that led to Eq. 1 gives

$$r_{XY,Y} = \mathrm{P}(x \equiv y, \mathrm{Y}) = \tfrac{1}{2}[\mathrm{P}(x \equiv y_m, \mathrm{Y}) + \mathrm{P}(x \equiv y_p, \mathrm{Y})] \tag{5}$$

and substituting Eqs. 4a and 4b into Eq. 5 yields

$$r_{XY,Y} = r_{XM,Y} + r_{XP,Y} - \mathrm{P}(x = y, \mathrm{Y})\,[\tfrac{1}{2}(F_M + F_P) + F_Y]\ . \tag{6}$$

Because $\mathrm{P}(x = y, \mathrm{Y}) = r_{XY,Y}/(1 + F_Y)$ (see Appendix C), rearranging Eq. 6 yields

$$r_{XY,Y} = (r_{XM,Y} + r_{XP,Y})\,(1 + F_Y)/[1 + \tfrac{1}{2}(F_M + F_P) + 2F_Y]\ . \tag{7}$$

In a noninbred population ($F_M = F_P = F_Y = 0$), so that

$$r_{XY,Y} = r_{XM,Y} + r_{XP,Y}\ .$$

For the last term in Eq. 3a, $\bar{\mathrm{Y}}$ does not contain the path from M to X through Y. Therefore, the random gene y_m is no closer than m to X, so that $\mathrm{P}(x \equiv y_m, \bar{\mathrm{Y}})$ is equal to $\mathrm{P}(x \equiv m, \bar{\mathrm{Y}})$. Proof that

$$\mathrm{P}(x \equiv y_m, \bar{\mathrm{Y}}) = \mathrm{P}(x \equiv m, \bar{\mathrm{Y}}) = r_{XM,\bar{Y}} \tag{8a}$$

is given in Appendix B.

Thus, if genes are identical by descent only through $\bar{\mathrm{Y}}$, then the probability that x is identical by descent to y_m is equal to the probability that x is identical by descent to m. If X is not a direct descendent of Y, genes can be identical by descent only through $\bar{\mathrm{Y}}$. Therefore, from Eq. 3a,

$$\mathrm{P}(x \equiv y_m) = \mathrm{P}(x \equiv y_m, \bar{\mathrm{Y}})$$

and similarly,

$$\mathrm{P}(x \equiv m) = \mathrm{P}(x \equiv m, \bar{\mathrm{Y}}) \ .$$

Using the above, together with Eq. 8 a, gives

$$\mathrm{P}(x \equiv y_m) = \mathrm{P}(x \equiv y_m, \bar{\mathrm{Y}}) = \mathrm{P}(x \equiv m, \bar{\mathrm{Y}}) = \mathrm{P}(x \equiv m) = r_{\mathrm{XM}} \ . \tag{8 a'}$$

Similarly for the last probability in Eq. 3 b, that x is identical by descent to the paternal gene in $\mathrm{Y}(y_p)$, it can be seen that

$$\mathrm{P}(x \equiv y_p, \bar{\mathrm{Y}}) = \mathrm{P}(x \equiv p, \bar{\mathrm{Y}}) = r_{\mathrm{XP}, \bar{\mathrm{Y}}} \ . \tag{8 b}$$

Also, if X is not a direct descendent of Y, x and p can be identical by descent only through $\bar{\mathrm{Y}}$. Therefore,

$$\mathrm{P}(x \equiv y_p) = \mathrm{P}(x \equiv y_p, \bar{\mathrm{Y}}) = \mathrm{P}(x \equiv p, \bar{\mathrm{Y}}) = \mathrm{P}(x \equiv p) = r_{\mathrm{XP}} \tag{8 b'}$$

Now, the same reasoning that led to Eq. 1 gives

$$r_{\mathrm{XY}, \bar{\mathrm{Y}}} = \mathrm{P}(x \equiv y, \bar{\mathrm{Y}}) = \tfrac{1}{2}[\mathrm{P}(x \equiv y_m, \bar{\mathrm{Y}}) + \mathrm{P}(x \equiv y_p, \bar{\mathrm{Y}})] \tag{9}$$

so that, from Eqs. 8 a and 8 b, $r_{\mathrm{XY}, \bar{\mathrm{Y}}}$ in Eq. 9 may be expressed as

$$r_{\mathrm{XY}, \bar{\mathrm{Y}}} = \tfrac{1}{2}[r_{\mathrm{XM}, \bar{\mathrm{Y}}} + r_{\mathrm{XP}, \bar{\mathrm{Y}}}] \tag{10}$$

If X is not a direct descendent of Y, Eq. 10 may be written as Eq. 2, which was shown in Eqs. 8 a′ and 8 b′.

Substituting Eqs. 7 and 10 into Eq. 3 yields the general expression for the coancestry between X and Y:

$$\begin{aligned} r_{\mathrm{XY}} = {} & (r_{\mathrm{XM}, \mathrm{Y}} + r_{\mathrm{XP}, \mathrm{Y}})\,(1 + F_{\mathrm{Y}})/[1 + \tfrac{1}{2}(F_{\mathrm{M}} + F_{\mathrm{P}}) + 2\,F_{\mathrm{Y}}] \\ & + \tfrac{1}{2}[r_{\mathrm{XM}, \bar{\mathrm{Y}}} + r_{\mathrm{XP}, \bar{\mathrm{Y}}}] \end{aligned} \tag{11}$$

Equation 11 does not require the condition that X is not a direct descendent of Y. If X is a direct descendent of Y, then $(r_{\mathrm{XM}, \mathrm{Y}} + r_{\mathrm{XP}, \mathrm{Y}}) > 0$, and Eq. 2 will yield an incorrect value for coancestry between X and Y. If X is not a direct descendent of Y, then $(r_{\mathrm{XM}, \mathrm{Y}} + r_{\mathrm{XP}, \mathrm{Y}}) = 0$ and $(r_{\mathrm{XM}, \bar{\mathrm{Y}}} + r_{\mathrm{XP}, \bar{\mathrm{Y}}}) = (r_{\mathrm{XM}} + r_{\mathrm{XP}})$, so that Eq. 2 will yield the correct value for conancestry.

In a noninbred population,

$$r_{\mathrm{XY}} = r_{\mathrm{XM}, \mathrm{Y}} + r_{\mathrm{XP}, \mathrm{Y}} + \tfrac{1}{2}(r_{\mathrm{XM}, \bar{\mathrm{Y}}} + r_{\mathrm{XP}, \bar{\mathrm{Y}}}) \ . \tag{12}$$

It also can be shown that, *if each of two individuals is not a direct descendent of the other*, coancestry between these two individuals may be computed as the average of the four coancestries between the parents of the two individuals. Let X and Y be relatives, where each is not a direct descendent of the other; and let A and B be parents of X, and C and D be parents of Y. Coancestry between X and Y may be computed as the average coancestry between X and the parents of Y,

$$r_{\mathrm{XY}} = \tfrac{1}{2}[r_{\mathrm{XC}} + r_{\mathrm{XD}}] \tag{13}$$

because X is not a direct descendent of Y. Alternatively, it may be computed as the average coancestry between Y and the parents of X,

$$r_{\mathrm{XY}} = \tfrac{1}{2}[r_{\mathrm{YA}} + r_{\mathrm{YB}}] \tag{14}$$

because Y is not a direct descendent of X.

For the same reason, C and D cannot be direct descendents of X. Therefore, coancestry between X and C may be computed as the average coancestry between C and the parents of X,

$$r_{\mathrm{XC}} = \tfrac{1}{2}[r_{\mathrm{AC}} + r_{\mathrm{BC}}] \ . \tag{15}$$

Similarly, coancestry between X and D is

$$r_{\mathrm{XD}} = \tfrac{1}{2}[r_{\mathrm{AD}} + r_{\mathrm{BD}}] \ . \tag{16}$$

Because X is not a direct descendent of Y, it follows that A and B cannot be direct descendents of Y. Therefore, coancestry between Y and A may be computed as the average coancestry between A and the parents of Y,

$$r_{\mathrm{YA}} = \tfrac{1}{2}[r_{\mathrm{AC}} + r_{\mathrm{AD}}] \ . \tag{17}$$

Similarly, coancestry between Y and B is

$$r_{\mathrm{YB}} = \tfrac{1}{2}[r_{\mathrm{BC}} + r_{\mathrm{BD}}] \ . \tag{18}$$

Substituting Eqs. 15 and 16 into Eq. 13, or Eq. 17 and 18 into Eq. 14, yields

$$r_{\mathrm{XY}} = \tfrac{1}{4}[r_{\mathrm{AC}} + r_{\mathrm{AD}} + r_{\mathrm{BC}} + r_{\mathrm{BD}}] \ , \tag{19}$$

a well known result (p. 89, Falconer 1989).

Numerical examples

Consider an example (Fig. 2) where X = F is a direct descendent of Y = E, such as parent-offspring, and where maternal parent M = A and paternal parent P = D, for Y, in a random mating population. Coancestry between X and Y by Eq. 1 is

$$r_{\mathrm{XY}} = \tfrac{1}{2}[\mathrm{P}(x \equiv y_m) + \mathrm{P}(x \equiv y_p)] \ .$$

Gene x can be identical by descent to y_m or to y_p if X inherits a copy of y_m or of y_p from Y (Y). Therefore, we may not substitute r_{XM} for $\mathrm{P}(x \equiv y_m)$ or r_{XP} for $\mathrm{P}(x \equiv y_p)$. Instead, upon appropriate substitutions from Eqs. 3 a, 4 a, 8 a, and C8, the first term in Eq. 1 is

$$\begin{aligned} \mathrm{P}(x \equiv y_m) &= 2\,r_{\mathrm{XM}, \mathrm{Y}} - r_{\mathrm{XY}, \mathrm{Y}}\,[F_{\mathrm{M}} + F_{\mathrm{Y}}]/(1 + F_{\mathrm{Y}}) + r_{\mathrm{XM}, \bar{\mathrm{Y}}} \\ &= 2(3/16) - (5/16)\,[0 + \tfrac{1}{4}]/(1 + \tfrac{1}{4}) + 0 \\ &= 5/16 \ , \end{aligned} \tag{20}$$

where from Eq. 3: $r_{\mathrm{XM}, \mathrm{Y}} = r_{\mathrm{XM}} - r_{\mathrm{XM}, \bar{\mathrm{Y}}} = 3/16 - 0 = 3/16$, because X is related to M only through the two paths that include the path in which X inherits genes from Y (Y); $r_{\mathrm{XY}, \mathrm{Y}} = r_{\mathrm{XY}} - r_{\mathrm{XY}, \bar{\mathrm{Y}}} = (3/8) - (1/16) = 5/16$, where $r_{\mathrm{XY}, \bar{\mathrm{Y}}}$ is computed from the pedigree after removing the path (Y) from Y to X in the calculations of the coancestry r_{XY}; and $F_{\mathrm{Y}} = \tfrac{1}{4}$.

From Eqs. 3 b, 4b, 8 b, and C8, the second term in Eq. 1 is

$$\begin{aligned} \mathrm{P}(x \equiv y_p) &= 2\,r_{\mathrm{XP}, \mathrm{Y}} - r_{\mathrm{XY}, \mathrm{Y}}\,[F_{\mathrm{P}} + F_{\mathrm{Y}}]/(1 + F_{\mathrm{Y}}) + r_{\mathrm{XP}, \bar{\mathrm{Y}}} \\ &= 2(3/16) - (5/16)[0 + \tfrac{1}{4}]/(1 + \tfrac{1}{4}) + 1/8 \\ &= 7/16 \ , \end{aligned} \tag{21}$$

where from Eq. 3: $r_{\mathrm{XP,Y}} = r_{\mathrm{XP}} - r_{\mathrm{XP},\bar{\mathrm{Y}}} = 5/16 - 1/8 = 3/16$, where $r_{\mathrm{XP},\bar{\mathrm{Y}}}$ is computed from the pedigree after removing the path from Y to X in the calculations of the coancestry r_{XP}.

Thus, by Eqs. 1, 20, and 21, the coancestry between X and Y is

$$r_{\mathrm{XY}} = \tfrac{1}{2}[(5/16) + (7/16)] = 3/8 \ ,$$

whereas

$$\begin{aligned}\tfrac{1}{2}[r_{\mathrm{XM}} + r_{\mathrm{XP}}] &= \tfrac{1}{2}[(3/16) + (5/16)]\\ &= \tfrac{1}{4} \neq r_{\mathrm{XY}}\end{aligned}$$

because X is a direct descendent of Y.

Consider another example (Fig. 2) where X = E is not a direct descendent of Y = C, such as uncle-nephew, where maternal parent M = A and paternal parent P = B, in a random mating population. Now,

$$\begin{aligned}\mathrm{P}(x \equiv y_m) &= 2r_{\mathrm{XM,Y}} - r_{\mathrm{XY,Y}}[F_{\mathrm{M}} + F_{\mathrm{Y}}]/(1 + F_{\mathrm{Y}}) + r_{\mathrm{XM},\bar{\mathrm{Y}}}\\ &= 0 + 3/8 = 3/8 \qquad (22)\end{aligned}$$

because X is not a direct descendent of Y; therefore, there are no paths through which X inherits genes from Y. Similarly,

$$\begin{aligned}\mathrm{P}(x \equiv y_p) &= 2r_{\mathrm{XP,Y}} - r_{\mathrm{XY,Y}}[F_{\mathrm{P}} + F_{\mathrm{Y}}]/(1 + F_{\mathrm{Y}}) + r_{\mathrm{XP},\bar{\mathrm{Y}}}\\ &= 0 + 1/8 = 1/8 \qquad (23)\end{aligned}$$

Thus, by Eqs. 1, 22, and 23, the coancestry between X and Y is

$$r_{\mathrm{XY}} = \tfrac{1}{2}[(3/8) + (1/8)] = \tfrac{1}{4}\,.$$

Now,

$$\begin{aligned}\tfrac{1}{2}[r_{\mathrm{XM}} + r_{\mathrm{XP}}] &= \tfrac{1}{2}[(3/8) + (1/8)]\\ &= \tfrac{1}{4} = r_{\mathrm{XY}}\end{aligned}$$

because X is not a direct descendent of Y.

Finally, because each of X and Y is not a direct descendent of the other, coancestry between X = E and Y = C may be computed as the average of the four coancestries between their parents, A and B, and A and D (Fig. 2):

$$\begin{aligned}r_{\mathrm{XY}} &= \tfrac{1}{4}[r_{\mathrm{AA}} + r_{\mathrm{AD}} + r_{\mathrm{BA}} + r_{\mathrm{BD}}]\\ &= \tfrac{1}{4}[\tfrac{1}{2} + \tfrac{1}{4} + 0 + \tfrac{1}{4}] = \tfrac{1}{4}\end{aligned}$$

because A is noninbred, and A and B are unrelated.

Discussion

We have provided an explanation for the condition under which coancestry between two individuals may be computed as the average coancestry between one individual and the parent of the other. The condition is that the former individual is not a direct descendent of the latter. Thus, for relatives where each individual is not a direct descendent of the other, coancestry between them may be computed as the average coancestry between one and the parents of the other or vice versa. For relatives where one individual is a direct descendent of the other, coancestry between them may not be computed as the average coancestry between the younger and the parents of the older.

In the tabular method, this condition is never violated because coancestry between an individual in one generation and another individual in a previous generation is never computed as the average coancestry between the former individual and the parents of the latter. Coancestry for parent-offspring, for example, is never computed as the average coancestry between the offspring and its grandparents.

Coancestry between an individual in one generation and another individual in a previous generation, if the former is not a direct descendent of the latter, may be computed as the average coancestry between the former individual and the parents of the latter. Coancestry between uncle and nephew, for example, may be computed as the average coancestry between the nephew and the parents of the uncle. In the tabular method, however, this coancestry is always computed as the average coancestry between the uncle and the parents of the nephew. The two computations give the same answer.

We have shown also that, if each of two individuals is not a direct descendent of the other, coancestry between these two individuals may be computed also as the average of the four conancestries between the parents of the two individuals. This condition, that each individual cannot be a direct descendent of the other, is not always stated unambiguously.

The method of computing coancestry as the average of the four coancestries may be used to compute coancestry between relatives of the same generation (Falconer 1989) or between relatives of different generations, if each individual is not a direct descendent of the other, as shown here and illustrated by the second numerical example.

Appendix

A *Proof that* $\mathrm{P}(x \equiv y_m, \mathrm{Y}) = 2r_{\mathrm{XM,Y}} - \mathrm{P}(x = y, \mathrm{Y})\,[F_{\mathrm{M}} + F_{\mathrm{Y}}]$

Remember that y_m is a random gene from M, and that it can be the maternal gene ($m_m = y_m$) or the paternal gene ($m_p = y_m$), each with probability $\tfrac{1}{2}$. Thus, expanding the first term in Eq. 3a yields

$$\begin{aligned}\mathrm{P}(x \equiv y_m, \mathrm{Y}) &= \mathrm{P}(x \equiv m_m, m_m = y_m, \mathrm{Y}) + \mathrm{P}(x \equiv m_p, m_p = y_m, \mathrm{Y})\\ &= \mathrm{P}(x \equiv m_m, \mathrm{Y}|m_m = y_m)\mathrm{P}(m_m = y_m)\\ &\quad + P(x \equiv m_p, \mathrm{Y}|m_p = y_m)\,\mathrm{P}(m_p = y_m)\\ &= \tfrac{1}{2}[\mathrm{P}(x \equiv m_m, \mathrm{Y}|m_m = y_m) + \mathrm{P}(x \equiv m_p, \mathrm{Y}|m_p = y_m)]\,. \qquad (\mathrm{A}\,1)\end{aligned}$$

We are considering that x can be identical by descent to y_m because X inherited a copy of y_m from Y(Y). Therefore, the fact that the random gene y_m is m_m now alters the probability that x is identical by descent to m_m, so that $\mathrm{P}(x \equiv m_m, \mathrm{Y}|m_m = y_m) \neq P(x \equiv m_m, \mathrm{Y})$, as explained below.

The random gene x can be identical by descent to m_m, with y_m being m_m or with y_m being m_p. Thus, the probability that x is identical by descent to m_m can be expanded as

$$\begin{aligned}P(x\equiv m_m, Y) &= P(x\equiv m_m, m_m=y_m, Y)+P(x\equiv m_m, m_p=y_m, Y)\\ &= P(x\equiv m_m, Y|m_m=y_m)\,P(m_m=y_m)\\ &\quad + P(x\equiv m_m, Y|m_p=y_m)\,P(m_p=y_m)\\ &= \tfrac{1}{2}[P(x\equiv m_m, Y|m_m=y_m)+P(x\equiv m_m, Y|m_p=y_m)]\,. \end{aligned} \tag{A 2}$$

Rearranging Eq. A 2 yields

$$P(x\equiv m_m, Y|m_m=y_m)=2\,P(x\equiv m_m, Y)-P(x\equiv m_m, Y|m_p=y_m)\ . \tag{A 3}$$

Consider the last term in Eq. A 3. The random gene x can be identical by descent to m_m through Y, given that Y receives m_p from M, in one of four ways:

x is $x_m (x=x_m)$, x_m is $y_m (x_m=y_m)$, and m_p is identical by descent to $m_m (m_p\equiv m_m)$;

or

x is $x_m (x=x_m)$, x_m is $y_p (x_m=y_p)$, and y_p is identical by descent to $m_m (y_p\equiv m_m)$;

or

x is $x_p (x=x_p)$, x_p is $y_m (x_p=y_m)$, and m_p is identical by descent to $m_m (m_p\equiv m_m)$;

or

x is $x_p (x=x_p)$, x_p is $y_p (x_p=y_p)$, and y_p is identical by descent to $m_m (y_p\equiv m_m)$ (Fig. 1).

Therefore, the last term in Eq. A 3 can be written as

$$\begin{aligned}&P(x\equiv m_m, Y|m_p=y_m)\\ &\quad = P(x=x_m, x_m=y_m, m_p\equiv m_m, Y)\\ &\qquad + P(x=x_m, x_m=y_p, y_p\equiv m_m, Y)\\ &\qquad + P(x=x_p, x_p=y_m, m_p\equiv m_m, Y)\\ &\qquad + P(x=x_p, x_p=y_p, y_p\equiv m_m, Y)\\ &\quad = \tfrac{1}{2}\{[P(x_m=y_m, Y)+P(x_p=y_m, Y)]\,F_M\\ &\qquad + [P(x_m=y_p, Y)+P(x_p=y_p, Y)]\,P(m_m\equiv y_p, Y)\}\\ &\quad = \tfrac{1}{2}[2\,P(x=y_m, Y)\,F_M+2\,P(x=y_p, Y)\,P(m_m\equiv y_p, Y)]\\ &\quad = P(x=y, Y)\,[F_M+P(m_m\equiv y_p, Y)]\end{aligned}$$

because $P(x=y_m, Y)=P(x=y_p, Y)=P(x=y, Y)$; F_M, the coefficient of inbreeding for M, is the probability that m_m is identical by descent to m_p. Then Eq. A 3 can be written as

$$\begin{aligned}&P(x\equiv m_m, Y|m_m=y_m)\\ &\quad = 2\,P(x\equiv m_m, Y)-P(x=y, Y)\,[F_M+P(m_m\equiv y_p, Y)].\end{aligned} \tag{A 4}$$

From the above, it is clear that

$$P(x\equiv m_m, Y|m_m=y_m)\not\equiv P(x\equiv m_m, Y).$$

Similarly for m_p

$$\begin{aligned}&P(x\equiv m_p, Y|m_p=y_m)\\ &\quad = 2\,P(x\equiv m_p, Y)-P(x=y, Y)\,[F_M+P(m_p\equiv y_p, Y)\end{aligned} \tag{A 5}$$

and $P(x\equiv m_p, Y|m_p=y_m)\not\equiv P(x\equiv m_p, Y)$.

Substituting Eqs. A 4 and A 5 into Eq. A 1 yields

$$\begin{aligned}P(x\equiv y_m, Y) &= \tfrac{1}{2}[2\,P(x\equiv m_m, Y)+2\,P(x\equiv m_p, Y)]\\ &\quad - P(x=y, Y)\,[2F_M+P(m_m\equiv y_p, Y)+P(m_p\equiv y_p, Y)]\\ &= 2\,r_{XM,Y}-P(x=y, Y)\,[F_M+F_Y]\end{aligned} \tag{A 6}$$

because $r_{XM,Y}=\tfrac{1}{2}[P(x\equiv m_m, Y)+P(x\equiv m_p, Y)]$ and $F_Y=\tfrac{1}{2}[P(m_m\equiv y_p, Y)+P(m_p\equiv y_p, Y)]\ =P(y_m\equiv y_p, Y)$.

B *Proof that* $P(x\equiv y_m, \bar{Y})=P(x\equiv m, \bar{Y})=r_{XM,\bar{Y}}$

Consider the last term in Eq. 3 a, the probability that x is identical by descent to the maternal gene of Y (y_m). Remember that y_m is a random gene from M, and that it can be the maternal gene $(m_m=y_m)$ or the paternal gene $(m_p=y_m)$, each with probability $\tfrac{1}{2}$. Thus, expanding the last term in Eq. 3 a yields

$$\begin{aligned}P(x\equiv y_m, \bar{Y}) &= P(x\equiv m_m, m_m=y_m, \bar{Y})+P(x\equiv m_p, m_p=y_m, \bar{Y})\\ &= P(x\equiv m_m, \bar{Y}|m_m=y_m)\,P(m_m=y_m)\\ &\quad + P(x\equiv m_p, \bar{Y}|m_p=y_m)\,P(m_p=y_m)\\ &= \tfrac{1}{2}[P(x\equiv m_m, \bar{Y}|m_m=y_m)+P(x\equiv m_p, \bar{Y}|m_p=y_m)]\,.\end{aligned} \tag{B 1}$$

There are two mutually exclusive sets of paths between M and X: the set of paths that includes Y, which contains the path from M to Y, and the set of remaining paths. For example, let Y = D, X = F, and M = A (Fig. 2). The first set comprises the path A → D → E → F, which contains the path D → E → F, and the second set comprises the remaining path A → E → F.

Sampling of genes from M to Y is independent of sampling of genes in the second set of paths. Thus, sampling of y_m from M in the first set does not affect $P(x\equiv m_m, \bar{Y})$, so that

$$P(x\equiv m_m, \bar{Y}|m_m=y_m)=P(x\equiv m_m, \bar{Y})$$

and, similarly for m_p,

$$P(x\equiv m_p, \bar{Y}|m_p=y_m)=P(x\equiv m_p, \bar{Y})\,.$$

Now Eq. B 1 can be written as

$$\begin{aligned}P(x\equiv y_m, \bar{Y}) &= \tfrac{1}{2}[P(x\equiv m_m, \bar{Y})+P(x\equiv m_p, \bar{Y})]\\ &= P(x\equiv m, \bar{Y})=r_{XM,\bar{Y}}\,.\end{aligned} \tag{B 2}$$

C *Proof that* $P(x=y, Y)=r_{XY,Y}/(1+F_Y)$

By definition,

$$r_{XY,Y}=P(x\equiv y, Y)\ . \tag{C 1}$$

Because x can be either x_m or x_p with probability $\tfrac{1}{2}$, and y can be either y_m or y_p with probability $\tfrac{1}{2}$, Eq. C 1 can be written as

$$\begin{aligned}r_{XY,Y} &= \tfrac{1}{4}[P(x_m\equiv y_m, Y)+P(x_m\equiv y_p, Y)\\ &\quad + P(x_p\equiv y_m, Y)+P(x_p\equiv y_p, Y)]\,.\end{aligned} \tag{C 2}$$

Consider the first term in Eq. C 2. The gene x_m can be identical by descent to y_m through Y in one of two ways:

x_m is $y_m (x_m=y_m)$ and y_m is identical by descent to $y_m (y_m\equiv y_m)$

or

x_m is $y_p (x_m=y_p)$ and y_m is identical by descent to $y_p (y_m\equiv y_p)$.

Therefore the first term in Eq. C 2 can be written as

$$\begin{aligned}P(x_m\equiv y_m, Y) &= P(x_m=y_m, Y)+P(x_m=y_p, Y)\,F_Y\\ &= P(x_m=y, Y)\,(1+F_Y)\end{aligned} \tag{C 3}$$

because $P(x_m=y_m, Y)=P(x_m=y_p, Y)=P(x_m=y, Y)$.

Similarly for the remaining terms in Eq. C 2,

$$P(x_m\equiv y_p, Y)=P(x_m=y, Y)\,(1+F_Y) \tag{C 4}$$

$$P(x_p\equiv y_m, Y)=P(x_p=y, Y)\,(1+F_Y) \tag{C 5}$$

$$P(x_p\equiv y_p, Y)=P(x_p=y, Y)\,(1+F_Y)\,. \tag{C 6}$$

Substituting Eqs. C 3 through C 6 into Eq. C 2 yields

$$\begin{aligned}r_{XY,Y} &= \tfrac{1}{2}[P(x_m=y, Y)+P(x_p=y, Y)]\,(1+F_Y)\\ &= P(x=y)\,(1+F_Y)\,.\end{aligned} \tag{C 7}$$

Rearranging Eq. C 7 yields

$$P(x=y, Y)=r_{XY,Y}/(1+F_Y)\,. \tag{C 8}$$

Acknowledgements. We thank one of the reviewers for suggestions that lead to a more accurate manuscript.

References

Doolittle DP (1987) Population genetics. Springer, Berlin Heidelberg New York

Emik LO, Terrill CE (1949) Systematic procedures for calculating inbreeding coefficients. J Hered 40:51–55

Falconer DS (1989) Introduction to quantitative genetics, 3rd edn. Longman, New York

Fernando RL, Grossman M (1990) Genetic evaluation with autosomal and X-chromosomal inheritance. Theor Appl Genet 80:75–80

Henderson CR (1976) A simple method for computing the inverse of a numerator relationship matrix in prediction of breeding values. Biometrics 32:69–83

Kempthorne O (1969) An introduction to genetic statistics. Iowa State University Press, Ames/IA

Malécot G (1969) The mathematics of heredity. Freeman, San Francisco/CA

Pirchner F (1983) Population genetics in animal breeding, 2nd edn. Plenum, New York

Van Vleck LD, Pollak EJ, Oltenacu EAB (1987) Genetics for the animal sciences. Freeman, New York