

Measuring Relatedness between Inbred Individuals

PHILIP W. HEDRICK AND ROBERT C. LACY

School of Life Sciences, Arizona State University, Tempe, AZ 85287 (Hedrick); and Chicago Zoological Society, Brookfield, IL 60513 (Lacy).

Address correspondence to Phil Hedrick at the address above, or e-mail: philip.hedrick@asu.edu.

Abstract

Genetic relatedness between individuals is an important measure in many areas of biology. However, some relatedness measures for use with molecular (allele) data assume that the individuals themselves are not inbred. Here, we present a new measure of relatedness based on the different modes of identity-by-descent for alleles that has an upper bound of 1 even when the individuals being compared are themselves inbred. This new measure is compared to several other measures of relatedness using several simple examples and pedigree data from the wolf population in Isle Royale National Park.

Subject areas: Conservation genetics and biodiversity

Key words: coefficient of relationship, inbreeding, pedigree, wolves

Measuring genetic relatedness between individuals within a population is widely carried out in research in evolution, behavior, and conservation. Methods to calculate relatedness from known pedigrees have long been available (Wright 1922; Henderson 1975) and are widely applied in animal and plant breeding (Falconer and Mackay 1996). More recently, a number of different approaches have been developed for estimating the amount of relatedness using molecular data on genotypes (Lynch and Ritland 1999; Blouin 2003; Weir et al. 2006; Wang 2014). As a working definition here, genetic relatedness is the proportion of homologous alleles shared between 2 individuals due to identity-by-descent from a common ancestor. Some of the situations in which measures of relatedness are used include estimating heritabilities in wild populations, estimating gene flow, maintaining diversity in managed populations, examining the genetic benefits of altruistic behavior, determining match probabilities for DNA profiles, and mapping disease genes (Blouin 2003; Weir et al. 2006).

For noninbred individuals, relatedness r_{xy} between 2 individuals x and y is equal to $2f_{xy}$ where f_{xy} is the coefficient of kinship between them. The coefficient of kinship is the probability that 2 alleles sampled from the same locus in the 2 individuals will be identical by descent and it is also equal to the inbreeding coefficient for a potential offspring from individuals x and y (Pamilo 1989). However, when individuals x and y are themselves inbred, the assumption of $r_{xy} = 2f_{xy}$ is not appropriate because both r_{xy} and f_{xy} have upper bounds of 1. For example, when $f_{xy} = 1$ (completely inbred parents and offspring),

then r_{xy} would be equal to 2, which clearly cannot be a valid measure of the proportion of alleles shared. To circumvent this problem, Wang (2011) defined his relatedness measure as equal to f_{xy} .

Often small populations that are of interest in evolution, behavior, or conservation have inbred individuals, an attribute that should be incorporated in measures of relatedness. Here we propose a new measure of relatedness applicable to genotypic data generated from DNA assays or simulations of allele transmission through pedigrees. This new measure has an upper bound of 1 and individuals always have $r = 1$ to themselves, even with inbreeding. We compare this new measure to the coefficient of relationship measure of Wright (1922) using data from several simple hypothetical examples and the complex pedigree for the wolf population from Isle Royale National Park.

Measures of Relatedness between Inbred Individuals

Inbreeding can be calculated using the different modes of identity-by-descent of 4 homologous genes possessed by 2 individuals (Jaquard 1972; Lynch and Walsh 1998). As given in Table 1, the 2 alleles in individual x are indicated as a and b and the 2 homologous alleles in individual y are indicated as c and d . For example, if there is a probability that x is inbred, then alleles a and b might be identical-by-descent. Or, if individuals x and y are related, they might share alleles that would be identical-by-descent in an offspring from them.

Table 1 The 9 identity states (Jaquard 1972) for the 2 alleles in 2 diploid individuals where states identical by descent are connected by a line

Identity state	Probability	Contribution to f_{xy}	Contribution to r_{xy}
	Δ_1	1	1
	Δ_2	0	0
	Δ_3	1/2	3/4
	Δ_4	0	0
	Δ_5	1/2	3/4
	Δ_6	0	0
	Δ_7	1/2	1
	Δ_8	1/4	1/2
	Δ_9	0	0

The 2 alleles in individual x are a and b and the 2 alleles in individual y are c and d . Horizontal lines indicate homozygosity in an individual from identity by descent. For the given identity states, the Δ_i values give the probability of identity-by-descent and the right 2 columns indicate the contribution to the inbreeding coefficient f_{xy} of an offspring from the individuals x and y and relatedness r_{xy} between individuals x and y .

Given the probability Δ_i of the 9 different identity-by-descent modes (Table 1), the inbreeding coefficient of an offspring between individuals x and y is

$$f_{xy} = \Delta_1 + \frac{1}{2}(\Delta_3 + \Delta_5 + \Delta_7) + \frac{1}{4}\Delta_8 \quad (1a)$$

This value measures the probability of identity-by-descent of alleles of a potential offspring from individuals x and y . If there is no past inbreeding, then only Δ_7 and Δ_8 can be non-zero and

$$f_{xy} = \frac{1}{2}\Delta_7 + \frac{1}{4}\Delta_8 \quad (1b)$$

Below we will use expression (1a) when there is past inbreeding.

Given that individuals x and y are potentially inbred (either alleles a and b in individual x or alleles c and d in individual y are identical-by-descent), then sharing of alleles between individuals x and y occurs for the 5 identity states 1, 3, 5, 7, and 8. For identity states 1 and 7, both alleles are shared between

the 2 individuals. Note the difference for identity state 7 where for relatedness, it is assumed that there is complete sharing (1) of alleles identical by descent unlike for inbreeding where the probability of identity-by-descent in an offspring is only 1/2 for identity state 7. Similarly for identity state 8, it is assumed that there is 1/2 sharing of alleles identical by descent for relatedness compared to that for inbreeding where there is a 1/4 probability of identity-by-descent in an offspring. The 2-fold difference in the contributions of identity states 7 and 8 to the inbreeding coefficient and relatedness occurs because identity-by-descent measured for inbreeding (f_{xy}) is assessing a property of the offspring and includes a generation of segregation from the parental individuals in which relatedness (r_{xy}) is measured.

For identity states 3 and 5, the proportion of alleles shared identical-by-descent are based on the following logic. For identity state 3 (similar logic can be used for identity state 5), let us assume that individual x has genotype A_1A_1 and that individual y has genotype A_1A_2 . First, individual x shares all (1) of its alleles with individual y (it does not have any alleles that y does not have, it only has allele A_1 which individual y also has). Second, individual y shares 1/2 of its alleles with x (A_1 is shared and A_2 is not shared). Hence weighting these 2 comparisons equally, the average sharing of alleles identical-by-descent of individuals x and y is $0.5(1) + 0.5(1/2) = 3/4$. In other words in a behavioral sense, individual y would not detect any alleles in individual x that are different from what it has itself while individual x would detect in individual y 1 allele that is shared and one that is not shared. Similarly, both Lynch (1988) and Li et al. (1993) reached the same conclusion about the 3/4 level of sharing for these identity states for analogous DNA fingerprint genotypes.

Therefore, the overall relatedness when there is inbreeding in 1 or both of the individuals, measured as the shared proportion of homologous alleles identical-by-descent, is

$$r_{xy} = \Delta_1 + \Delta_7 + \frac{3}{4}(\Delta_3 + \Delta_5) + \frac{1}{2}\Delta_8 \quad (2a)$$

If there is no past inbreeding, then only Δ_7 and Δ_8 can be non-zero and using expression equation (1b)

$$r_{xy} = \Delta_7 + \frac{1}{2}\Delta_8 = 2f_{xy} \quad (2b)$$

Another measure, known as the coefficient of relationship (Crow and Kimura 1970, p. 69), that takes into account past inbreeding is from Wright (1922)

$$r_{xy}^* = \frac{2f_{xy}}{[(1 + f_x)(1 + f_y)]^{1/2}} \quad (3)$$

where f_x and f_y are the inbreeding coefficients in individuals x and y and f_{xy} is the expected inbreeding coefficient of an offspring from individuals x and y . This measure was developed for calculating inbreeding from the correlation of additive genetic effects between the dam and sire (described initially by Wright as the correlation between uniting gametes) by solving expression (3) for f_{xy} . This correlation coefficient has been used to measure relatedness between individuals, and (as will be shown below) does not give the same value as the proportion of alleles identical-by-descent shared between 2 individuals.

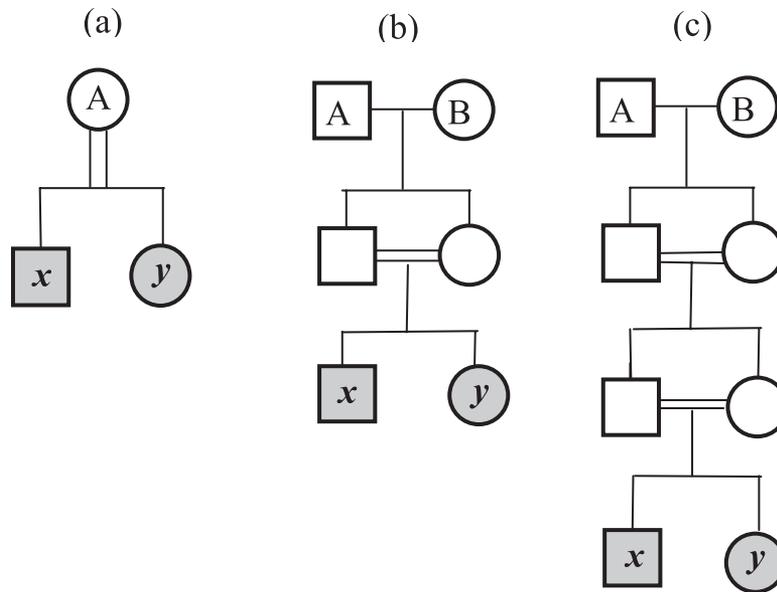


Figure 1. Pedigrees for the 3 hypothetical examples with inbred individuals discussed in the text and Table 2 (a) 2 sibs from self-fertilization of a parental individual A (indicated by the 2 parallel vertical lines), (b) sibs whose parents are sibs, and (c) sibs whose parents and grandparents are sibs. For (b) and (c) the 2 ancestors are A and B and these are full sibs from 1 and 2 generations of full-sib mating, respectively.

Table 2 Three simple hypothetical examples in which relatedness is calculated for a pair of inbred individuals

Probability	Two-selfed sibs	Sibs whose parents are sibs	Sibs whose parents and grandparents are sibs
Δ_1	0.125	0.0625	0.17188
Δ_2	0.125	0.03125	0.02734
Δ_3	0.25	0.125	0.15625
Δ_4	0	0.03125	0.01953
Δ_5	0.25	0.125	0.15625
Δ_6	0	0.03125	0.01953
Δ_7	0.25	0.21875	0.25391
Δ_8	0	0.3125	0.17969
Δ_9	0	0.0625	0.01562
$f_x = f_y$	0.5	0.25	0.375
f_{xy}	0.5	0.375	0.5
$2f_{xy}$	1.0	0.75	1.0
r_{xy}	0.75	0.625	0.75
r_{xy}^*	0.667	0.6	0.727

The values of Δ_i for 2 selfed sibs (2 sibs from self-fertilization of a parental individual) are from Wang (2011) and Δ_i values for sibs whose parents are sibs and sibs whose parents and grandparents are sibs (full sibs from 1 and 2 generations of full-sib mating) are from Karigl (1981).

Some Hypothetical Examples

To illustrate the application of this new relatedness measure and compare it to other measures, we calculated values for 3 simple hypothetical examples with inbred individuals (Figure 1). The first example (a) assumes that there are 2 siblings (x and y) that have been produced by self-fertilization from a single parental individual (Wang 2011). The second (b) and third (c) examples assume that there are 2 full siblings (x and y) that have been produced by either 1 or 2 generations of full-sib mating (Karigl 1981).

Table 2 gives the 9 Δ_i values for these 3 examples and the resulting f_{xy} values which are 0.5, 0.375, and 0.5, respectively.

Clearly, relatedness cannot be equal to $2f_{xy}$ here because $2f_{xy}$ is equal to 1.0 in both the first and third examples and they do not share all their alleles. On the other hand, for the 3 examples the relatedness estimate is $r_{xy} = 0.75, 0.625,$ and 0.75 , greater than f_{xy} but much less than unity. For the 3 examples, the coefficient of relationship $r_{xy}^* = 0.667, 0.6,$ and 0.727 , which in all examples is similar in magnitude to, but slightly less than, r_{xy} .

Isle Royale Wolf Pedigree

The wolf population on Isle Royale National Park in Lake Superior has been relatively small (50 or less) since its natural founding by migrants in the late 1940s. In 1997, a male wolf

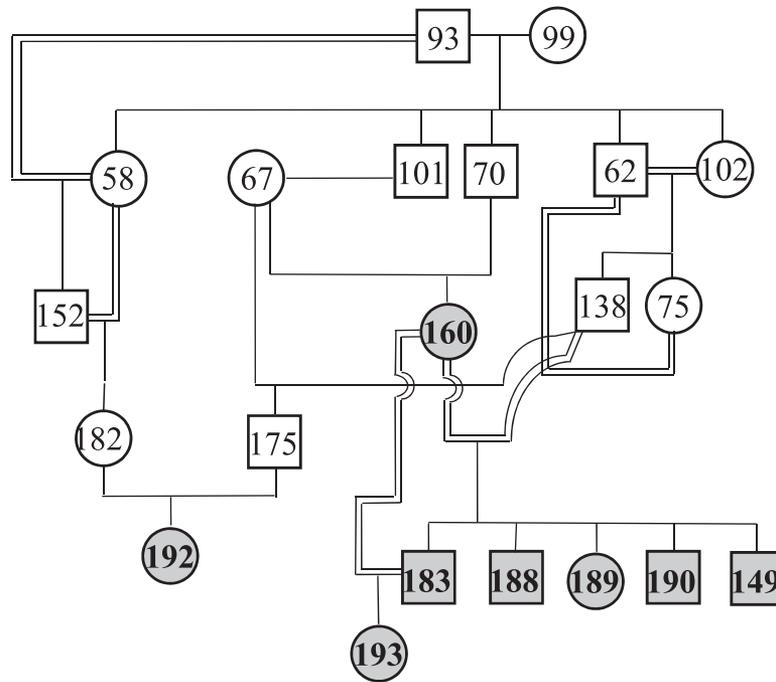


Figure 2. The pedigree of Isle Royale wolves giving the known ancestors of the 8 wolves genetically identified (shaded) on the island in 2012. All of the wolves are descended from M93 who migrated to the island in 1997 and 2 females, F99 his first mate, and F67, another unrelated resident female. Here squares represent males, circles represent females, and double lines indicate first-degree matings, either between a parent and an offspring or between 2 full sibs.

M93 (M indicates male and F indicates female) immigrated into the population across an ice bridge and his ancestry quickly dominated the population (Adams et al. 2011; Hedrick et al. 2014). In 2012, there were only 9 wolves observed and in 2013 only 8 wolves were observed (only 8 wolves in 2012 and 7 wolves in 2013 were identified genetically). In 2012, there was no reproduction for the first time since the population has been monitored. Figure 2 gives the pedigree established from genetic analysis for the 8 genetically identified wolves in 2012 and their ancestors (Adams et al. 2011; Hedrick et al. 2014). As a result, to understand the constitution of potential mating pairs, it was important to estimate the relatedness between these potential mates, the 4 males and the 4 females, all of whom were inbred except for 1 female F160.

Table 3 gives the proportion of the 5 identity states observed that contribute to relatedness between the potential mates for the 5 different possible matings. Three of the males (M149, M188, and M190) are full sibs with an inbreeding coefficient of 0.125 and have the same relationship with the females in population. The fourth full-sib male, M183, is the father of 1 of the females, F193. Overall, the relatedness was the lowest for the mating with 1 of the 4 full-sib males and F192 because the males are first-order relatives (sons, siblings, or parent) of the other 3 living females. F160 was the only potential female parent that was not inbred because she was the offspring of 2 unrelated individuals, M70 and F67. As a result, for the mating with her as maternal parent, both Δ_1 and Δ_5 are zero and Δ_3 , Δ_7 , and Δ_8 , which give mother-son transmission, sum to 1. F189 is a full sib of the 4

male wolves. As a result, for the mating with her as maternal parent, Δ_7 and Δ_8 , where there is sharing of parental alleles but no homozygosity, are the highest for this mating.

Overall, the relatedness between potential parents is quite high for F160, F189, and F193 ($r_{xy} = 0.595$, and 0.617 , and 0.567 , respectively) and the male full sibs. The relatedness for the potential parents M183 and F193 (mother and son), is even higher at 0.734 . All these values are significantly higher than for F192 and the male full sibs for which $r_{xy} = 0.396$. This observation that a mating of 1 of the full-sib males and F192 has a much lower relatedness is especially relevant because there appears to have been some reproduction in 2013 and it is thought to have been from a mating between M190 and F192. In other words, a potential mating with lowest relatedness of the 5 possible combinations appears to have been successful.

As expected, all the values of r_{xy} are less than $2f_{xy}$. The largest difference is for the potential mating between M183 and F193 where $r_{xy} = 0.734$ and $2f_{xy} = 0.876$ (83.8%) and the smallest difference is for the matings of M149, M183, M188, or M190 and F160 where $r_{xy} = 0.595$ and $2f_{xy} = 0.624$ (95.4%).

The 2 measures of relatedness, r_{xy} and r_{xy}^* , are very similar for the potential matings, even with inbreeding in the pedigree. For the 3 potential matings with lower maternal inbreeding (F160, F189, and F192), r_{xy}^* averages about 1% less than r_{xy} while for the potential matings with F193 in which there is higher maternal inbreeding r_{xy}^* is 4.8% less than r_{xy} for the matings with M149, M188, and M190 and 1.9% less than r_{xy} for the mating with M183.

Table 3 The 5 identity states that contribute to relatedness due to alleles identical-by-descent between individuals and their probabilities Δ_i for the potential matings on Isle Royale in 2012

Identity state	Probability	M149, M183, M188, M190			M149, M188, M190	M183
		F160	F189	F192	F193	F193
	Δ_1	0	0.040	0.010	0.035	0.064
	Δ_3	0.125	0.061	0.049	0.062	0.065
	Δ_5	0	0.062	0.076	0.156	0.255
	Δ_7	0.127	0.273	0.065	0.169	0.241
	Δ_8	0.748	0.425	0.454	0.399	0.376
	f_y	0.000	0.125	0.125	0.312	0.312
	f_{xy}	0.312	0.344	0.219	0.328	0.438
	r_{xy}	0.595	0.617	0.396	0.567	0.734
	r_{xy}^*	0.589	0.611	0.389	0.540	0.720

The 4 potential male mates are the top individual x with alleles a and b and these 4 males are all full sibs with an inbreeding coefficient of 0.125. The 4 males are equivalent for F160, F189, and F192 (all with the values given in those 3 columns), while F193 is more closely related to M183 than to the other 3 males. At the bottom are the inbreeding coefficients for the female f_p for an offspring from the female and one of the full-sib males or M183 f_{xy} , and the relatedness, measured by r_{xy} and r_{xy}^* , between the female and one of the full-sib males or M183.

Discussion

A new measure of relatedness r_{xy} based on the different modes of identity-by-descent given here allows the individuals to be themselves inbred and has a range from 0 to 1. For many populations in which there has been previous inbreeding, this measure is appropriate. Although our new measure of relatedness and the coefficient of relationship of Wright (1922) give very similar results, they are not identical because they are estimating slightly different concepts. The measure of Wright (1922) gives the correlation of additive genetic effects between the 2 individuals, while r_{xy} measures the quantity of interest in many studies, that is, the proportion of homologous alleles shared by 2 individuals identical-by-descent.

Both the new relatedness measure r_{xy} and the measure of Wright r_{xy}^* are also similar to twice the coefficient of kinship (f_{xy}), but the 3 values diverge when there is inbreeding in the pedigree. For many purposes, such as when calculating inbreeding coefficients or estimating the loss of gene diversity across generations of a pedigreed population (Lacy 1995), f_{xy} is the appropriate measure as it directly translates into the reduction of heterozygosity. However, for some purposes, the proportion of shared alleles between 2 individuals is the property of interest. For example, kin selection theory is based on r_{xy} , although in the derivation of inclusive fitness, Hamilton (1964) originally noted that the appropriate measure of genetic overlap between 2 individuals will be equal to Wright's coefficient of relatedness only in the absence of inbreeding, and Hamilton's description of relatedness only encompassed identity states Δ_7 and Δ_8 as in equation (2b). The importance of using r_{xy} rather than $2f_{xy}$ in considerations of kin selection

can be seen in that 2 heterozygous individuals that share genotypes (Δ_7) are considered less closely related than are homozygous individuals that share genotypes (Δ_1) with respect to $2f_{xy}$ (but not with r_{xy}), which would be not match the concept of relatedness in Hamilton's kin selection theory.

Another important application of r_{xy} is the estimation of variance components from resemblance among relatives. The additive genetic correlation among relatives is given by Wright's measure, r_{xy}^* , but its calculation (expression 3) requires knowledge of the inbreeding coefficients, derived from f_{xy} values in the prior generation and therefore requiring additional pedigree information. Often, methods such as the Animal Model use $2f_{xy}$ as the coefficient for the additive genetic covariance between relatives, with the stated assumption that there is no inbreeding in the pedigree (Lynch and Walsh 1998). However, even when inbreeding cannot be assumed not to exist within the studied populations, the additive genetic covariance between relatives can be estimated as the percent of shared alleles, r_{xy} .

Our measure of relatedness should be useful in understanding the estimated level of relatedness from genomic data that use runs of homozygosity (ROH) (Pemberton et al. 2012) along chromosomal segments and identity of chromosomal segments between individuals x and y . With such complete data, the components of the relatedness measure ($\Delta_1, \Delta_3, \Delta_5, \Delta_7$, and Δ_8) could be estimated separately (and even potentially compared to that expected from pedigree information).

The new measure of relatedness from proportions of shared alleles should be useful in many small populations that have a history of inbreeding. In such populations, this measure could help understand and predict the evolutionary,

conservation, and behavioral consequences of relatedness of individuals that are, or are not, inbred.

Funding

Ullman Professorship for partial funding (P.W.H.).

Acknowledgments

We thank J.A. Vucetich, R.O. Peterson, and L.M. Vucetich for sharing data and insight that form the basis for the Isle Royale wolf example described in this paper. We also thank E. Thompson and J. Ballou for their insights into the topics discussed here.

References

- Adams JR, Vucetich LM, Hedrick PW, Peterson RO, Vucetich JA. 2011. Genomic sweep and potential genetic rescue during limiting environmental conditions in an isolated wolf population. *Proc Roy Soc B*. 278:3336–3344.
- Blouin MS. 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends Ecol Evol* 18:503–511.
- Crow JF, Kimura M. 1970. *An introduction to population genetics theory*. New York: Harper and Row.
- Falconer DS, Mackay TFC. 1996. *Introduction to quantitative genetics*, 4th edn. Harlow, UK: Longman.
- Hamilton WD. 1964. The genetical evolution of social behaviour. I. *J Theor Biol*. 7:1–16.
- Hedrick PW, Vucetich LM, Adams JR, Peterson RO, Vucetich JA. 2014. Genetic rescue: genetic analysis and the collapse of the Isle Royale wolf population. *Cons Genet* 15:1111–1121.
- Henderson CR. 1975. Rapid method for computing the inverse of a relationship matrix. *J Dairy Sci* 58:1727–1730.
- Jaquard A. 1972. Genetic information given by a relative. *Biometrics* 28:1101–1114.
- Karigl G. 1981. A recursive algorithm for the calculation of identity coefficients. *Ann Hum Genet*. 45:299–305.
- Lacy RC. 1995. Clarification of genetic terms and their use in the management of captive populations. *Zoo Biol* 14:565–577.
- Li CC, Weeks DE, Chakravarti A. 1993. Similarity of DNA fingerprints due to chance and relatedness. *Hum Hered*. 43:45–52.
- Lynch M. 1988. Estimation of relatedness by DNA fingerprinting. *Mol Biol Evol*. 5:584–599.
- Lynch M, Ritland K. 1999. Estimation of pairwise relatedness with molecular markers. *Genetics*. 152:1753–1766.
- Lynch M, Walsh B. 1998. *Genetics and analysis of quantitative traits*. Sunderland (MA): Sinauer Associates.
- Pamilo P. 1989. Comparison of relatedness estimators. *Evolution* 44:1378–1382.
- Pemberton TJ, Absher D, Feldman MW, Myers RM, Rosenberg NA, Li JZ. 2012. Genomic patterns of homozygosity in worldwide human populations. *Am J Hum Genet*. 91:275–292.
- Wang J. 2011. COANCESTRY: a program for simulating, estimating and analysing relatedness and inbreeding coefficients. *Mol Ecol Resour*. 11:141–145.
- Wang J. 2014. Marker-based estimates of relatedness and inbreeding coefficients: an assessment of current methods. *J Evol Biol* 27:518–530.
- Weir BS, Anderson AD, Hepler AB. 2006. Genetic relatedness analysis: modern data and new challenges. *Nat Rev Genet*. 7:771–780.
- Wright S. 1922. Coefficients of inbreeding and relationship. *Amer Natur* 56:330–338.

Received April 10, 2014; First decision October 6, 2014;
Accepted October 7, 2014

Corresponding Editor: C. Scott Baker