A. F. Groen · B. W. Kennedy · J. J. Eissen

Potential bias in inbreeding depression estimates when using pedigree relationships to assess the degree of homozygosity for loci under selection

Received: 26 January 1995 / Accepted: 3 February 1995

Abstract A potential bias in estimation of inbreeding depression when using pedigree relationships to assess the degree of homozygosity for loci under selection is indicated. A comparison of inbreeding coefficients based on either pedigree or genotypic frequencies indicated that, as a result of selection, the inbreeding coefficient based on pedigree might not correspond with the random drift of allelic frequencies. Apparent differences in average levels of both inbreeding coefficients were obtained depending on the genetic model (additive versus dominance, initial allelic frequencies, heritability) and the selection system assumed (no versus mass selection). In the absence of selection, allelic frequencies within a small population change over generations due to random drift, and the pedigree-based inbreeding coefficient gives a proper assessment of the accompanying probability of increased homozygosity within a 'replicate' by indicating the variance of allelic frequencies over replicates. With selection, in addition to random drift, directional change in allelic frequencies is not accounted for by the pedigree-based inbreeding coefficient. This result implies that estimation of inbreeding depression for traits under either direct or indirect selection, estimated by a regression of performance on pedigree-based coefficients, should be carefully interpreted.

Key words Inbreeding · Finite-locus model · Selection

Introduction

Two individuals of common ancestry may each carry replicates of genes present in the ancestor and, when mated, they may pass to their offspring genes that are 'identical by descent'. Mating related individuals is called 'inbreeding'. The inbreeding coefficient of an individual is the average probability that two genes at any given locus are identical by descent (Falconer 1989). The average inbreeding coefficient of all individuals in a population expresses the amount of drift in allelic frequencies from a defined base population (Falconer 1989). As a result, inbreeding causes a change in the frequency of heterozygous animals and, therefore, a change (usually considered a reduction) in the mean for traits subject to positive dominance gene interactions: inbreeding depression. A broad survey of the effects of inbreeding on various performance traits in livestock is given by Pirchner (1985).

Models used to estimate inbreeding depression usually include inbreeding as a regression variable. For a model with \( n_l \) unlinked, biallelic loci in gametic-phase equilibrium, the theoretical value of the regression coefficient for inbreeding (Kemphorne 1957; De Boer and Van Arendonk 1992) equals:

\[
b = -2 \sum_{k=1}^{n_l} p_k q_k d_k
\]

where \( p \) and \( q \) are allelic frequencies and \( d \) is the genotypic value of the heterozygote at the \( k \)th locus. Early estimates of \( b \) were made on an intra-sire basis to account for some of the effects of both genetic and environmental time trends (e.g. Von Krosigk and Lush 1958; Allaire and Henderson 1965). More recently, mixed models have been applied, which correct for environmental trends by considering them as fixed effects (e.g. herd-year-season). Genetic trends are accounted for by random additive genetic effects, including additive genetic relationships (which account for in-
breeding; Casanova et al. 1992; Miglior et al. 1992), dominance effects, and additive*additive interaction effects (without accounting for inbreeding in the dominance and additive*additive relationship matrices; Miglior et al. 1993). This latter model was shown to yield seemingly unbiased estimates of additive genetic effects (breeding values) in the presence of dominance and inbreeding for both selected and unselected populations (Uimari and Kennedy 1990; De Boer and Van Arendonk 1992; De Boer and Hoeschele 1993; Johansson et al. 1993). However, using stochastic simulation with a finite number of loci, Uimari and Kennedy (1990) and De Boer and Van Arendonk (1992) showed that for populations under selection these models might give estimates of inbreeding depression that do not correspond to theoretical values. This may be because theoretical values assume gametic-phase equilibrium (De Boer and Van Arendonk 1992). Ignoring inbreeding in the dominance-relationship matrix and ignoring covariance between additive and dominance effects also can lead to biased estimates of the dominance effects (Johansson et al. 1993). A third reason might be the inaccuracy of pedigree relationships for determining inbreeding coefficients in the case of loci with alleles under selection. To quote Falconer (1989; p. 85): "When the coefficient of inbreeding is deduced from the pedigrees of real populations, it does not necessarily describe the state of dispersion of the gene frequencies. It is essentially a statement about the pedigree relationships, and its correspondence with the state of dispersion is dependent on the absence of the processes that counteract dispersion, in particular on selection being negligible."

The aim of the present study was to illustrate potential bias in inbreeding depression estimates when using pedigree relationships to assess the degree of homozygosity for loci under selection. The simulation study employed, using a finite-locus model, compared average inbreeding coefficients based on actual genotypic frequencies and those based on pedigree relationships.

Methods

Simulation model

The model simulated genotypic and phenotypic values for a single trait. An animal's genotypic value was the sum of its genetic values for 100 unlinked, biallelic loci. Allelic effects were (1/2) a and (-1/2) a. The genotypic value of the heterozygote was d. Thus additive genetic variance is:

$$\sigma^2_A = \sum_{l=1}^{100} 2pq[a + d(q - p)]^2$$

(1)

where p and q are frequencies of favorable and unfavorable alleles, respectively, in the first generation. A normally distributed environmental effect with standard deviation $\sigma_e$ was added to each genotypic value such that the narrow-sense heritability ($h^2$) was 0.30 in the first generation. Each simulation included nine generations. The first contained 75 males and 75 females whose genotypes were randomly generated assuming Hardy-Weinberg proportions and gametic-phase equilibrium, giving initial allelic frequencies (p and q). Genotypes of animals in later generations were simulated from parental genotypes. Selection, either randomly or on basis of own phenotype (mass selection), was both in males and females. The number of selected males ($N_m$), the number of selected females ($N_f$), with multiple females mated per male but only one male per female mated, and the number of progeny per mating ($N_p$), were to maintain a constant population size (150) over generations. The effect of inbreeding level was studied by varying the number of males and females selected. The probability of an offspring to be male or female was 50%.

Inbreeding coefficients

If inbreeding is the only source of disequilibrium, the inbreeding coefficient ($f$) in a population for a given locus can be estimated from the deviation between the expected ($\xi$) and observed ($\xi$) number of heterozygotes (Cune-Cohen 1982)

$$f = (\xi - \xi) / \xi$$

(2)

In a biallelic model, let the observed numbers of genotypes in a sample of n be $a_{11}$, $a_{12}$ and $a_{22}$ for the favorable homozygote, the heterozygote and the unfavorable homozygote, respectively, then

$$\xi = 2pqn, \quad p = 2a_{12}$$

(3)

where

$$p = (a_{11} + a_{12})/n, \quad q = (a_{12} + a_{22})/n$$

(4)

Equations (2), (3) and (4) can be combined to obtain

$$f = (a_{11}a_{22} - a_{12}a_{21})/(a_{11} + a_{12})(a_{12} + a_{22})$$

(5)

When information on genotypic frequencies is known for multiple unlinked loci, an averaged $f$ over loci can be derived (e.g. Bumstead et al. 1987). When the expected number of heterozygotes, $\xi$, is based on genotypic frequencies from the same generation for which f is estimated, equation (5) is expected to give non-zero inbreeding only when inbreeding is practiced as a mating system in the previous generation. With random mating in populations of finite size, inbreeding in terms of allelic dispersion can be assessed by using the allelic frequencies of the first generation to estimate expected genotypic frequencies in subsequent generations. In the present study, allelic frequencies in the first generation (for each alternative and replicate sample separately) were used to calculate the expected number of heterozygotes in subsequent generations.

Inbreeding coefficients on the basis of pedigree relationships were computed using the algorithm of Tier (1990). As a reference, a simple prediction of inbreeding trend was considered where $N_p$ is the effective population size and $f$ (averaged over animals) in generation t equals (Falconer 1989)

$$f_t = \frac{1}{2N_e} + \left(1 - \frac{1}{2N_e}\right)f_{t-1}$$

(6)

Alternatives

First, estimated inbreeding coefficients on the basis of genotypic frequencies and pedigree, and the predicted inbreeding coefficient on the basis of effective population size, were compared for a situation with no selection and random mating (assuming $a = 20$, $d = 0$, $h^2 = 0.30$, initial frequency $p = 0.5$) for different effective population sizes $N_e$: 100 ($N_{m} = 50$; $N_{f,m} = 1$; $N_f = 6$), 40 (15; 2; 10), and 20 (6; 5; 10). (N.B. derivation of $N_e$ from $N_m$, $N_f$, $N_{m,f}$, $N_f$ is correct for situations without selection, but is used for situations with selection as well.)

Second, random and phenotypic selection were compared at $N_e = 40$ (initial frequency $p = 0.5$ and $h^2 = 0.3$) under an additive model ($a = 20$ and $d = 0$), partial dominance model (20; 10), a complete dominance model (20; 20) and an overdominance model (20; 80).