GENETIC EVALUATION AND SELECTION USING GENOTYPIC, PHENOTYPIC AND PEDIGREE INFORMATION

Rohan L. Fernando

Department of Animal Science, Iowa State University, Ames IA 50011, USA

SUMMARY

Genetic evaluation is simplest when genotypes are known at all QTL. Then, the BLUE of genetic differences can be obtained by modeling phenotypic values using a fixed linear model. Genetic evaluation is also straightforward when genotypic information is available only for some of the QTL and no genotypic information is available for the remaining QTL. Then, BLUP of genetic differences can be obtained by modeling phenotypic values using a mixed linear model. It is expected that genotypes at markers linked to QTL will be available for a considerable period before genotypes at QTL will be available for genetic evaluation. For this situation, BLUP of genetic differences can be obtained by using a mixed linear model that includes gametic effects at the marked QTL.

Keywords genetic evaluation, genetic markers, marker assisted selection

INTRODUCTION

Genetic variability for most economically important traits is assumed to be due to the segregation of alleles at several quantitative trait loci (QTL). Until recently, little information has been available on individual QTL, and genetic evaluation and selection have been based on phenotypic and pedigree information. Research in molecular genetics during the last decade, however, has resulted in the identification of a large number of polymorphic marker loci and a few candidate genes (Bishop et al. 1994; Archibald et al. 1995; Rothschild et al. 1996). Several experimental procedures and statistical techniques are being used to identify markers that are closely linked to QTL (Weller 1986; Weller and Fernando 1991; Haley and Knott 1992; Hoeschele 1993a; van Arendonk et al. 199413; Xu and Atchley 1995; Stricker et al. 1996; Uimari et al. 1996; Darvasi 1997; Gringola et al. 1997). It is expected that closely linked markers will be used to identify the QTL themselves.

It has been proposed that genotypic information at markers and QTL can be combined with phenotypic information to improve genetic evaluation and selection (Soller 1978; Smith and Webb 1981; Smith and Simpson 1986). Statistical methods for combining genotypic and phenotypic information across large pedigrees are discussed in this paper.

STATISTICAL METHODS

Assumptions and Notation. Consider a trait for which alleles are segregating at N

QTL. For simplicity, we assume two alleles per QTL and additive inheritance. Then, the vector of additive genotypic values for n animals can be written as

$$a = \sum_{i=1}^{N} a_i = \sum_{i=1}^{N} Q_i \alpha_i \quad (1)$$

where a_i is the vector of additive genetic values at locus i, α_i is 2 x 1 vector of additive effects for the two alleles at locus i, and Q_i is an n x 2 incidence matrix relating the two additive effects at locus i to the animals; the jth row of Q_i is $[2,0]$ if animal j is homozygous for the first allele at loci i, $[1, 1]$ if it is heterozygous, or $[0,2]$ if it is homozygous for the second allele. For convenience, let the QTL be numbered such that: $1, \ldots, k$ are loci for which genotypic information is available for the QTL, $k+1, \ldots, k+m$ are loci for which genotypic information is not available for the QTL but is available for linked markers, and $k + m + 1, \ldots, N$ are loci for which genotypic information is not available for the QTL nor for linked marker loci. Also, let $Q^K = [Q_1, \ldots, Q_k]$, $Q^M = [Q_{k+1}, \ldots, Q_{k+m}]$ and $Q^U = [Q_{k+m+1}, \ldots, Q_N]$, and α^K, α^M and α^U denote the vectors of additive effects that correspond to Q^K, Q^M and Q^U, respectively. Phenotypic values on animals can be modeled as

$$y = X\beta + ZQ\alpha + e \quad (2)$$

where β is a vector of fixed effects, X is an incidence matrix relating β to y, Z is an incidence matrix relating a to y, Q = $[Q^K, Q^M, Q^U]$, α is the vector of additive effects for all N loci, and e is a residual vector with E(e) = 0 and Var(e) = $I\sigma_e^2$.

Genetic evaluation. For most of the current evaluations, k = 0 and m = 0. Then, at each locus i, $E(a_i) = 0$ and $\mathrm{Var}(a_i) = A\sigma_{a_i}^2$, where A is the additive relationship matrix and $\sigma_{a_i}^2$ is the additive genetic variance for the ith QTL. Thus, E(a) = 0 and, if the QTL are in gametic equilibrium, $\mathrm{Var}(a) = A\sigma_a^2$, where $\sigma_a^2 = \sum \sigma_{a_i}^2$. For this situation, the best linear unbiased predictor (BLUP) of a is obtained by solving Henderson's mixed model equations (Henderson 1973).

Genetic evaluation is even more straightforward when k = N. Then, the incidence matrix Q is known, and, conditional on Q, (2) is a fixed linear model with fixed effects β and α. For this situation, best linear unbiased estimates (BLUE) of genetic differences between animals can be easily obtained using least squares theory (Searle 1971), provided that genetic differences are not confounded with other fixed effects in the model.

Further, dominance and epistasis can be accommodated by including within locus and between locus allelic interactions. The method of analysis would be the same for linked or unlinked QTL. In principle, the same method can be used whether loci are in equilibrium or not, but including loci that are highly dependent can cause numerical problems due to collinearity of the incidence matrix. With repeated cycles of selection, genetic

evaluations will be biased because selection will be across fixed effects (Henderson 1975). This bias is due to the estimation of fixed genotypic effects, and as information accumulates the bias would disappear. Although genetic evaluation is straightforward for this situation, developing selection rules that maximize long-term genetic progress may not be as straightforward.

Nejati-Javaremi et al. (1997) have considered an alternative approach for the above situation. In their approach, when QTL genotypes are known, genetic evaluations are obtained by solving the following mixed model equations:

$$\begin{bmatrix} \boldsymbol{X'X} & \boldsymbol{X'Z} \\ \boldsymbol{TZ'X} & \boldsymbol{TZ'Z} + \boldsymbol{I}\sigma_e^2/\sigma_a^2 \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\beta}} \\ \hat{\boldsymbol{a}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'y} \\ \boldsymbol{TZ'y} \end{bmatrix} \tag{3}$$

where $\boldsymbol{T}$ is the total allelic identity matrix. The total allelic identity between individuals $\boldsymbol{x}$ and $\boldsymbol{y}$ was defined as

$$T_{xy} = \frac{\sum_{i=1}^{N} T_{xy,i}}{N}$$

where $T_{xy,i}$ is the total allelic identity between $\boldsymbol{x}$ and $\boldsymbol{y}$ for locus $\boldsymbol{i}$. Their definition of total allelic identity between $\boldsymbol{x}$ and $\boldsymbol{y}$ for locus $\boldsymbol{i}$ is equivalent to twice the probability that a randomly picked allele from locus $\boldsymbol{i}$ in x is identical in state to a randomly picked allele from locus $\boldsymbol{i}$ in $\boldsymbol{y}$.

The above approach has a Bayesian interpretation. To see this, let $\boldsymbol{\beta}$ have a flat prior distribution and let $\boldsymbol{\alpha}$ have a normal prior distribution with null mean and variance $\boldsymbol{I}\sigma_\alpha^2$. The ratio of variance components $\sigma_e^2/\sigma_\alpha^2$ is assumed to be known. Further, note that the total allelic identity matrix can be expressed as $\boldsymbol{T} = \frac{\boldsymbol{QQ'}}{2N}$. *Now,*

$$\begin{aligned} \text{Var}(\boldsymbol{a}|\text{QTL genotypes}, \sigma_\alpha^2) &= \boldsymbol{Q}\text{Var}(\boldsymbol{\alpha})\boldsymbol{Q'} \\ &= \boldsymbol{QQ'}\sigma_\alpha^2 \\ &= \boldsymbol{T}2N\sigma_\alpha^2 \end{aligned} \tag{4}$$

This conditional variance (4), with $\sigma_\alpha^2 = \frac{\sigma_a^2}{2N}$, is implied by the mixed model equations (3). Thus, from a Bayesian point of view, $\hat{\boldsymbol{a}}$ from (3) is the posterior mean of of a. It would be more straightforward to first get the posterior mean of $\boldsymbol{\alpha}$ as the solution to:

$$\begin{bmatrix} \text{x ' x} & \boldsymbol{X'ZQ} \\ \boldsymbol{Q'Z'X} & \boldsymbol{Q'Z'ZQ} + \boldsymbol{I}\sigma_e^2/\sigma_\alpha^2 \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\beta}} \\ \hat{\boldsymbol{\alpha}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X'y} \\ \boldsymbol{QZ'y} \end{bmatrix} \tag{5}$$

and then obtain the posterior mean of a as $\hat{\boldsymbol{a}} = \boldsymbol{Q}\hat{\boldsymbol{\alpha}}$.

When $0 < \boldsymbol{k} <$ N and $\boldsymbol{m}$ = 0, model (2) becomes

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{ZQ}^K\boldsymbol{\alpha}^K + \boldsymbol{ZQ}^U\boldsymbol{\alpha}^U + \boldsymbol{e} \tag{6}$$

In this model, $\boldsymbol{\beta}, \boldsymbol{\alpha}^K$, and $\boldsymbol{\alpha}^U$ are fixed effects. Given the observed genotypes, $\boldsymbol{Q}^K$ is a fixed matrix but $\boldsymbol{Q}^U$ is random, and therefore, $\boldsymbol{u} = \boldsymbol{Q}^U\boldsymbol{\alpha}^U$ is random. Suppose QTL k+ I,... , N are in gametic equilibrium with each other and with QTL 1,. . . , $\boldsymbol{k}$. Then, E(u) = 0 and $\mathrm{Var}(\boldsymbol{u}) = \boldsymbol{A}\sigma_U^2$, where $\sigma_U^2 = \sum_{i=k+1}^{N} \sigma_{a_i}^2$. Thus, genetic evaluations can be obtained from Henderson's mixed model equations for (6), where $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}^K$ are fixed effects and $\boldsymbol{u} = \boldsymbol{Q}^U\boldsymbol{\alpha}^U$ is random (Kennedy et al. 1992).

Now consider the situation where genotypic information is missing for some of the first $\boldsymbol{k}$ loci, on some of the animals. Suppose, for example, genotype $i < k$ is missing for animal j. Then, the elements of row j of $\boldsymbol{Q}_i$ are not observed. The simplest approach for this situation would be to replace these unobserved row elements in $\boldsymbol{Q}_i$ by their expected values conditional on the observed genotypes at locus i, and proceed with the analysis for complete genotypic information (Clamp et al. 1992; Kinghorn and Kerr 1995). It is not difficult to show that genetic evaluations obtained by this approach are unbiased. Other properties are not known. More accurate evaluations can be obtained by combined segregation and linkage analysis (Stricker et al. 1996; Uimari et al. 1996). However, if genotypes are missing at many of the QTL, this approach may not be computationally feasible.

Next, we consider genetic evaluation when $m > 0$. First, let $m = 1$, and suppose genotypes are observed at one marker that is in equilibrium with QTL $k + 1$. Now, model (2) can be written as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{Q}^K\boldsymbol{\alpha}^K + \boldsymbol{Z}\boldsymbol{Q}^M\boldsymbol{\alpha}^M + \boldsymbol{Z}\boldsymbol{Q}^U\boldsymbol{\alpha}^U + \boldsymbol{e} \tag{7}$$

In this model, $\boldsymbol{\beta}, \boldsymbol{\alpha}^K, \boldsymbol{\alpha}^M$, and $\boldsymbol{\alpha}^U$ are fixed effects. But, the matrix $\boldsymbol{Q}^M$ is random, therefore, $\boldsymbol{w} = \boldsymbol{Q}^M\boldsymbol{\alpha}^M$ is random. Because the marker locus and QTL $\boldsymbol{k}$ + 1 are in equilibrium, $\mathrm{E}(\boldsymbol{w}|\boldsymbol{M}) = 0$, where $\boldsymbol{M}$ denotes the marker information. But, $\mathrm{Var}(\boldsymbol{w}|\boldsymbol{M}) \neq \boldsymbol{A}\sigma_{k+1}^2$, in general (Fernando and Grossman 1989). For efficient computations, it is more convenient to work with gametic values v than with additive genetic values w, and the model (7) can be written in terms of $\boldsymbol{u}$ and v as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{Q}^K\boldsymbol{\alpha}^K + \boldsymbol{Z}\boldsymbol{K}\boldsymbol{v} + \boldsymbol{Z}\boldsymbol{u} + \boldsymbol{e} \tag{8}$$

where $\boldsymbol{K}$ is an incidence matrix relating gametes to genotypic values, and w = $\boldsymbol{K}\boldsymbol{v}$. If parental origin of the marker alleles can be inferred from marker genotypes, recursive formulae given by Fernando and Grossman (1989) can be used to compute G = $\mathrm{Var}(\boldsymbol{v}|\boldsymbol{M})$ and its inverse. The matrix G is a function of the recombination rate between the marker and the linked QTL and of the additive variance at the QTL. These parameters can be estimated by maximum likelihood (Weller and Fernando 1991; van Arendonk et al. 1994b; Gringola et al. 1997).

Formulae have been developed to accommodate the situation where parental origin of the marker alleles is not known (Wang et al. 1991; Hoeschele 1993b; van Arendonk et al.

Table **1. Pedigree with missing marker genotype for animal 3**

Animal	Genotype	Sire	Dam
1	A_1A_2	0	0
2	A_1A_3	0	0
3	missing	1	2
4	A_1A_2	0	0
5	A_1A_2	3	4
6	A_1A_2	3	4

1994a; Wang et al. 1995). Let $\boldsymbol{G}_i$ be the covariance matrix for gametic values of animals $1, \ldots, i$. Then, the tabular method to compute G can be expressed in matrix notation as

$$\boldsymbol{G}_{i+1} = \begin{bmatrix} \boldsymbol{G}_i & \boldsymbol{G}_i\boldsymbol{L} \\ \boldsymbol{L}'\boldsymbol{G}_i & \boldsymbol{C}_{i+1} \end{bmatrix} \tag{9}$$

where $\boldsymbol{C}_{i+1}$ is a 2 x 2 matrix with $\sigma^2_{k+1}/2$ on the diagonal and $f_{i+1}\sigma^2_{k+1}/2$ on the off-diagonals, f_{i+1} is the conditional inbreeding coefficient for animal $i + 1$ given the marker information, and $\boldsymbol{L}$ is a $2i$ x 2 matrix (Wang et al. 1995). If the marker genotypes of the parents are not missing, each column of $\boldsymbol{L}$ will contain at most four non-zero elements that are located in rows corresponding to the gametes of the parents of animal $i+1$. This leads to an efficient method to invert G (van Arendonk et al. 1994a; Wang et al. 1995). Thus, genetic evaluations can be obtained from Henderson's mixed model equations for (8).

When genotypes of parents are missing, however, $\boldsymbol{L}$ may not have a simple form, and inverting G may not be simple. For example, consider the pedigree in Table **1,** where the genotype M_3 for animal 3 is missing. Then, G can be computed as

$$\boldsymbol{G} = \sum_{M_3} \boldsymbol{G}_{|M_3} \Pr(M_3|\boldsymbol{M}) \tag{10}$$

where the summation is over all possible values for M_3 and $\boldsymbol{G}_{|M_3}$ is the value of $\boldsymbol{G}$ computed with animal 3 having genotype M_3 (Hoeschele 1993b; Wang et al. 1995). When the marker genotypes are complete, $\boldsymbol{L}$ has a simple; known form and $\boldsymbol{G}_{|M_3}$ can be computed easily using (9). Once G is computed for a pedigree with missing genotypes using (10), one can solve for $\boldsymbol{L}$ of any animal. For the pedigree in Table 1, $\boldsymbol{L}$ for animal

6 is

$$L = \begin{vmatrix} 0.073 & -0.040 \\ -0.069 & 0.080 \\ 0.073 & -0.043 \\ -0.028 & -0.012 \\ 0.252 & 0.190 \\ 0.247 & 0.078 \\ 0.207 & 0.066 \\ 0.079 & 0.441 \\ 0.255 & -0.062 \\ -0.090 & 0.302 \end{vmatrix} \tag{11}$$

When $\boldsymbol{L}$ is dense as above, the inverse of G cannot be obtained efficiently, and further, $\boldsymbol{G}^{-1}$ will not be sparse. Then, obtaining genetic evaluations using Henderson's mixed model equations may not be feasible. It has been proposed to use a simple $\boldsymbol{L}$ matrix to approximate G when marker genotype information is not complete (Hoeschele 1993b; Wang et al. 1995). The consequences of such approximations need to be studied.

Now suppose marker information is available for m > 1 unlinked QTL. Genetic evaluations for this situation can be obtained by expanding model (8) to include the gametic effects of the m QTL with marker information. This will result in $2nm$ equations for gametic effects being included in the mixed model equations. An alternative is to combine the gametic effects for the m QTL with marker information and the u QTL without marker information as

$$\mathrm{c} = \sum_{i=k+1}^{k+m} \boldsymbol{K}\boldsymbol{v}_i + \boldsymbol{u} \tag{12}$$

where $\boldsymbol{v}_i$ is the gametic effect for the ith QTL, and the variance of c given the marker information is

$$\mathrm{Var}(\boldsymbol{c}|\boldsymbol{M}) = \boldsymbol{K} \sum_{i=k+1}^{k+m} \mathrm{Var}(\boldsymbol{v}_i|\boldsymbol{M})\boldsymbol{K}' + \mathrm{Var}(\boldsymbol{u}) \tag{13}$$

van Arendonk et al. (1994a) have proposed an algorithm to invert $\mathrm{Var}(\boldsymbol{c}|\boldsymbol{M})$. Their algorithm may be useful in very particular situations, but for a general pedigree with marker information on most animals, in comparisons that we have made, their method was less efficient than inversion based on the Cholesky decomposition. Further, although the inverse matrix of each $\mathrm{Var}(\boldsymbol{v}_i|\boldsymbol{M})$ and $\mathrm{Var}(\boldsymbol{u})$ is very sparse, the inverse of $\mathrm{Var}(\boldsymbol{c}|\boldsymbol{M})$ is a dense matrix. Thus, the mixed model equations for a model with c in it will be very dense (van Arendonk et al. 1994a).

Hoeschele (1993) showed that the inverse of $\mathrm{Var}\begin{bmatrix}\boldsymbol{c}|\boldsymbol{M}\\ \boldsymbol{v}|\boldsymbol{M}\end{bmatrix}$ is sparse and can be computed efficiently ($\mathrm{v}' = [\boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_{k+m}]$). Further, she showed that this efficiency is not lost by

eliminating from v: gametic effects of animals that do not have marker genotypes and are not in the pedigree path connecting genotyped descendants, or the gametic effect of a genotyped animal without offspring, if the gametic effect was inherited from an unknown parent or from a parent with both gametic effects eliminated from v. Thus, equations for certain gametic effects could be eliminated from the mixed model equations without making them dense. Reduced animal models can also be employed to reduce the number of equations (Cantet and Smith 1991; Saito and Iwaisaki 1996).

When some of the QTL with marker information are linked, Goddard (1992) proposed to model each $\boldsymbol{v}_i$ conditional on markers flanking the ith QTL. He gave formulae to compute $\mathrm{Var}(\boldsymbol{v}_i|\boldsymbol{M})$ and its inverse. These results are also based on a simple, known form for $\boldsymbol{L}$ in (9). When the linkage phase between the flanking markers is known, $\boldsymbol{L}$ has a simple, known form. Even with complete marker genotype information, the linkage phase of some parents may not be known and approximations will have to be employed.

Genetic evaluations can also be undertaken by computing the conditional mean of a given the genotypic, phenotypic, and pedigree information. This would require specifying completely the distribution of the random effects in the model and computing the mean by methods related to complex segregation and linkage analysis. Gibbs sampling can be used to reduce the computational burden. It has been shown, however, that the Gibbs sampler may not converge when more than two alleles are segregating at a locus.

ACKNOWLEDGMENTS

Discussions with Daniel Gianola, Liviu Totir and Bruce Southey are gratefully acknowledged. Research support was provided by award no. 94-37205-1307 of the National Research Initiative Competitive Grants Program of the USDA.

REFERENCES

Archibald, A. L., Haley, C. S., Brown, J. F., Couperwhite, S., McQueen, H. A., Nicholson, D., Coppieters, W., Van de Weghe, A., Stratil, A., Wintero, A. K. and et al. (1995) ***Mamm. Genome.*** 6:157–175.

Bishop, M. D., Kappes, S. M., Keele, J. W., Stone, R. T., Sunden, S. L. F., Hawkins, G. A., Toldo, S. S., Fries, R., Grosz, M. D., Yoo, J. and Beattie, C. W. (1994) ***Genetics*** 136:619–639.

Cantet, R. J. C. and Smith, C. (1991) ***Genet. Sel. Evol.*** 23:221–233.

Clamp, P. A., Beever, J. E., Fernando, R. L., McLaren, D. G. and Schook, L. B. (1992) ***J. Anim. Sci.*** 70:2695–2706.

Darvasi, A. (1997) ***Mamm. Genome.*** 8:163–167.

Fernando, R. L. and Grossman, M. (1989) ***Genet. Sel. Evol.*** 21:467–477.

Goddard, M. E. (1992) *Theor. Appl. Genet.* 83:878–886.

Gringola, F. E., Hoeschele, I. and Tier, B. (1997) *GSE* 28:479–490.

Haley, C. S. and Knott, S. A. (1992) ***Heredity*** 69:315–324.

Henderson, C. R. (1973) Proc. "Anim. Breed. Genet. Symp. in Honor of Dr. J. L. Lush", pp. 1041. Amer. Soc. Anim. Sci. and Amer. Dairy Sci. Assoc., Champaign, IL.

Henderson, C. R. (1975) ***Biometrics*** 31:423–449.

Hoeschele, I. (1993) *Theor. Appl. Genet.* 85:946–952.

Hoeschele, I. (1993) ***J. Dairy Sci.*** 76:1693–1713.

Kennedy, B. W., Quiton, M. and Van Arendonk, J. A. M. (1992) ***J. Anim. Sci.*** 70:2000–2012.

Kinghorn, B. P. and Kerr, R. J. (1995) Proc. "Proc. Eleventh Conf. Australian Assoc. of Anim. Breeding and Genetics". University of New England, Armidale, Australia.

Nejati-Javaremi, A., Smith, C. and Gibson, J. P. (1997) ***J. Anim. Sci.*** 75:1738–1745.

Rothschild, M., Jacobson, C., Vaske, D., Tuggle, C., Wang, L. Short, T., Eckardt, G., Sasaki, S., Vincent, A., McLaren, D., Southwood, O., van der Steen, H., Mileham, A. and Plastow, G. (1996) *Proc. Natl. Acad. Sci. USA* 93:201–205.

Saito, S. and Iwaisaki, H. (1996) ***Genet.*** *Sel. Evol.* 28:465–477.

Searle, S. R. (1971) Linear Models. John Wiley and Sons, New York.

Smith, C. and Simpson, ***S.*** P. (1986) ***J. Anim. Breed. Genet.*** 103:205–217.

Smith, C. and Webb, A. J. (1981) ***J. Anim. Breed. Genet.*** 98:161.

Soller, M. ***(1978)*** *Anim. Prod.* 27:133–139.

Stricker, C., Fernando, R. L. and Elston, R. C. (1996) ***Genetics*** 141:1651–1656.

Uimari, P., Thaller, G. and Hoeschele, I. (1996) ***Genetics*** 143:1831–1842.

van Arendonk, J. A. M., Tier, B. and Kinghorn, B. (1994) ***Genetics*** 137:319–329.

van Arendonk, J. A. M., Tier, B. and Kinghorn, B. P. (1994) Proc. "17th Int. Congr. Genet.", p. 192. Birmingham, UK.

Wang, T., Fernando, R. L., van der Beek, S., Grossman, M. and van Arendonk, J. A. M. (1995) ***Genet.*** *Sel. Evol.* 27:251–274.

Wang, T., van der Beek, S., Fernando, R. L. and Grossman, M. (1991) ***J. Anim. Sci.*** 69(Suppl. 1):202.

Weller, J. and Fernando, R. L. (1991) . In: "Gene Mapping: Strategies, Techniques and Applications", pp. 305-328, Editors L. B. Schook, H. A. Lewin and D. G. McLaren, Marcel Dekker.

Weller, J. I. (1986) ***Biometrics*** 42:627–640.

Xu, S. and Atchley, W. R. (1995) ***Genetics*** 141:1189–1197.