Maximum-Likelihood Estimation of Relatedness

Brook G. Milligan[1]

Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, Scotland

Manuscript received July 17, 2002
Accepted for publication November 26, 2002

ABSTRACT

Relatedness between individuals is central to many studies in genetics and population biology. A variety of estimators have been developed to enable molecular marker data to quantify relatedness. Despite this, no effort has been given to characterize the traditional maximum-likelihood estimator in relation to the remainder. This article quantifies its statistical performance under a range of biologically relevant sampling conditions. Under the same range of conditions, the statistical performance of five other commonly used estimators of relatedness is quantified. Comparison among these estimators indicates that the traditional maximum-likelihood estimator exhibits a lower standard error under essentially all conditions. Only for very large amounts of genetic information do most of the other estimators approach the likelihood estimator. However, the likelihood estimator is more biased than any of the others, especially when the amount of genetic information is low or the actual relationship being estimated is near the boundary of the parameter space. Even under these conditions, the amount of bias can be greatly reduced, potentially to biologically irrelevant levels, with suitable genetic sampling. Additionally, the likelihood estimator generally exhibits the lowest root mean-square error, an indication that the bias in fact is quite small. Alternative estimators restricted to yield only biologically interpretable estimates exhibit lower standard errors and greater bias than do unrestricted ones, but generally do not improve over the maximum-likelihood estimator and in some cases exhibit even greater bias. Although some nonlikelihood estimators exhibit better performance with respect to specific metrics under some conditions, none approach the high level of performance exhibited by the likelihood estimator across all conditions and all metrics of performance.

AN understanding of the relatedness between individuals plays an important role in many areas of population biology and genetics. For example, it is central to quantitative genetics and plays a crucial role in estimating heritability and additive genetic variances and covariances (Falconer 1981; Lynch and Walsh 1998). Likewise, it may be useful in studies of isolation-by-distance or population structure. Consequently, a number of different means of quantifying relatedness have been developed. Most inclusive of these are the sets of identity-by-descent modes described by Jacquard (1974). However, for large noninbred populations these reduce to a pair of quantities: the probability that two individuals share two alleles identical-by-descent and the probability that they share one allele identical-by-descent. More commonly, the coefficient of coancestry θ (Jacquard 1974) or the coefficient of relatedness $r = 2\theta$ are used to quantify the degree of relatedness between two individuals.

Estimates of θ or r may be derived in a variety of ways. Traditionally, they are calculated from a known pedigree (Crow and Kimura 1970). Increasingly, however, molecular marker data have been used to estimate relatedness. Consequently, a number of estimators have been developed for this purpose (Thompson 1975; Queller and Goodnight 1989; Li *et al.* 1993; Ritland 1996a; Lynch and Ritland 1999; Wang 2002). These have been developed in a variety of different ways. The Queller and Goodnight (1989) estimator was motivated to ensure that Hamilton's rule (Hamilton 1964a,b) applied under general circumstances given an estimate of relatedness r; in contrast, the Ritland (1996a), Lynch and Ritland (1999), and Wang (2002) estimators were based on different method-of-moments approaches to the relationship between relatedness and genotypic similarity.

Both Ritland (1996a) and Lynch and Ritland (1999) mention maximum-likelihood estimators of relatedness coefficients; however, they dismiss their utility on the basis of simulations that indicate that many (*e.g.*, 70 or more) loci may be required. However, both approaches deviate from the traditional approach involving a likelihood function defined for the set of three parameters sufficient for describing relatedness in a noninbred population (Thompson 1975). While the likelihood function used by Ritland (1996a, p. 180) is a special case applicable to a single two-gene relatedness parameter (*e.g.*, θ), that proposed by Lynch and Ritland (1999, Equation 12) cannot be derived from the traditional one (see appendix b). Further, both admitted solutions outside the biologically meaningful parameter space

[1] *Address for correspondence:* Department of Biology, New Mexico State University, Las Cruces, NM 88003. E-mail: brook@nmsu.edu.

(Thompson 1975) and thus yield estimates that cannot be interpreted as probabilities of identity-by-descent. Thus, despite extensive efforts to characterize the statistical behavior of other estimators of relatedness, it remains unclear how the traditional-likelihood one compares.

The statistical behavior of relatedness estimators is critical to their utility in practice. Because of their complexity, simulations have generally been relied upon to characterize the sampling error. These simulations have taken two approaches. The first constructed data sets from relatively simple conditions, assuming identical allele frequency distributions across loci (Ritland 1996a; Lynch and Ritland 1999). A second approach (Van de Casteele *et al.* 2001) was motivated by actual microsatellite data. In this case, data sets were constructed from more complex situations involving variation among loci in allele frequency distributions and variation in population structure. A similar approach involving actual data and a known pedigree was used to test the utility of marker-based estimators (Thomas *et al.* 2002). These two basic approaches have yielded somewhat contradictory pictures of the range of estimators available. It is clear, however, that not all estimators perform equally well under all conditions. Further, the same estimator may not perform best under all conditions. Which estimator performs best may depend not only on the nature of the biological conditions, but also on the criterion used to measure performance.

Conspicuously lacking from the array of relatedness estimators under test is the traditional general maximum-likelihood estimator (Thompson 1975). Often maximum-likelihood estimators exhibit many desirable features (Kendall *et al.* 1979), including having lower standard error, being asymptotically unbiased, being adaptable to a wide variety of sampling conditions, and naturally accounting for differences among different subsets of the sample, for example, different allele-frequency distributions among loci. The primary negative feature of maximum-likelihood estimators is that they are often biased for finite sample sizes. However, in many cases that bias is small enough to be biologically irrelevant or can be dramatically reduced with reasonable sampling.

Given the many desirable features of maximum-likelihood estimators, it would be natural to develop one for relatedness of individuals. Further, it would be useful to determine whether a maximum-likelihood estimator of relatedness approaches its asymptotic properties rapidly enough to be useful in practice or to compete with nonlikelihood estimators. In this study, we investigate one such estimator (Thompson 1975) on the basis of the genetic information available for two individuals assayed at many different loci. In addition to the genotypic information, the estimator relies on the allele frequency distributions at each locus sampled. For this analysis we assume that the allele-frequency distributions are known without error to focus on the behavior of the relatedness estimator itself. In practice, realistic samples will often involve enough individuals that errors in the allele-frequency distribution will be quite small. In some cases, for example, microsatellites segregating many alleles, errors in the frequencies may be significant. The importance of the additional sampling variance introduced is beyond the scope of this article.

Note that the approach taken here, to estimate relationship as a continuous parameter, is distinct from a related one, which infers the degree of relatedness from among a set of discrete possibilities. Both approaches are discussed by Thompson (1975). The latter approach is further developed by Thompson (1986) and has been used extensively in human genetics recently to classify groups of individuals (usually pairs) into distinct relationship classes (Boehnke and Cox 1997; Painter 1997; Broman and Weber 1998; Sieberts *et al.* 2002).

The primary goal of this study is to assess the performance of the likelihood estimator, in comparison with some of those already developed, under a range of biological sampling conditions. In particular, the performance is quantified by two measures of the distribution of estimates obtained for each estimator (the standard error and the bias) and by the overall deviation of those estimates from the parametric value, quantified by the root mean-square error. From this information it is possible to determine how aspects of the sampling conditions, *e.g.*, number of loci or segregating alleles, or aspects of the relationship being estimated influence the ability of one estimator or another to perform well. Thus, some guidance for experimental design can be developed.

The initial focus of this study is on a relatively simple set of conditions, although one meant to mimic a variety of natural situations. In this sense, it is more closely connected to the evaluation used by Lynch and Ritland (1999) than to that used by Van de Casteele *et al.* (2001). This approach has been chosen to provide a clearer indication of how each different factor, *e.g.*, number of loci sampled, number of segregating alleles, and allele frequencies, influences the performance of each estimator.

STATISTICAL MODELS

For population samples in which the ancestry of individual alleles is unknown, the most general means of describing the relatedness of one individual to another is in terms of the nine identity modes described by Jacquard (1974) (Figure 1). The degree of relatedness is quantified by a set of coefficients $\Delta = (\Delta_1, \Delta_2, \ldots \Delta_9)$, each of which represents the probability of the four alleles at a single locus in two diploid individuals sharing the corresponding particular pattern of identity-by-descent. In a large, noninbred population, only Δ_7, Δ_8, and Δ_9 are nonzero; consequently, in such a popula-

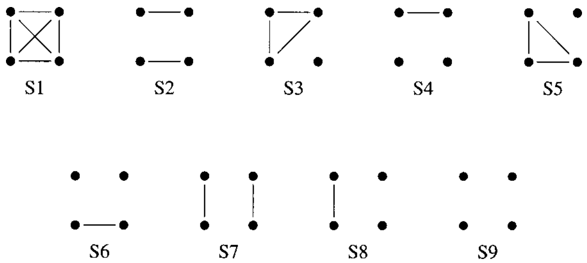

Figure 1.—Modes of identity-by-descent between two individuals. In each figure the two upper dots represent the two alleles in one individual, while the two lower dots represent the two alleles in the second individual. The lines indicate alleles that are identical-by-descent.

tion any pattern of relationship between two individuals can be described by that set of three coefficients. The most commonly used summary of the degree of relationship is the coefficient of coancestry (Jacquard 1974; Lynch and Walsh 1998),

$$\theta = \Delta_1 + \frac{1}{2}(\Delta_3 + \Delta_5 + \Delta_7) + \frac{1}{4}\Delta_8, \qquad (1)$$

which quantifies the probability that two individuals will produce an inbred offspring were they to mate. This latter coefficient plays central roles in the estimation of heritability and additive genetic variance (Falconer 1981; Lynch and Walsh 1998) and in the definition of inclusive fitness (Hamilton 1964a,b). Hence, estimators of relatedness (Queller and Goodnight 1989; Li *et al.* 1993; Ritland 1996a; Lynch and Ritland 1999; Wang 2002) primarily focus on θ or $r = 2\theta$ even though there is a loss of information in the process of transforming the complete set of parameters Δ into a single quantity θ. For development of a likelihood estimator of relatedness, the underlying complete set of parameters Δ is used; subsequently θ (or r) can be calculated from Equation 1 if necessary.

Likelihood models: Likelihood estimators are based on a probability model of the sampled data. In this case, the unit of sampling is a pair of individuals, each one of which has been assayed genetically at L loci. The estimator described here is based on the assumption of independently segregating marker loci. The likelihood for the overall sample, therefore, is simply the product of the likelihoods across the loci.

The basic probability model of the sampled alleles at a single diploid locus is well known (Thompson 1975). Usually it is given only for the case of large, noninbred populations where only three modes of relatedness (S_7, S_8, and S_9) are possible. However, the structure of the model is much clearer in its general form and it is applicable to the full range of population structures. As a basis of further generalization it thus warrants explicitly outlining the complete one-locus likelihood model for the nine relatedness coefficients.

There are nine distinct patterns of identity-in-state for the four alleles sampled at a single locus in two individuals. Table 1 lists these in the second column. As an example, a pair of individuals each homozygous for allele A_1 represent identity-in-state mode $\mathcal{S}_1$, whereas two individuals of genotypes A_1A_1 and A_1A_2 represent identity-in-state mode $\mathcal{S}_3$.

Given that a pair of individuals is known to be related according to identity-by-descent mode S_j, the probability of each identity-in-state pattern S_i is dependent on the allele frequencies. For example, if two noninbred individuals have two identical-by-descent alleles (S_7), either of two identity-in-state patterns ($\mathcal{S}_1$ or $\mathcal{S}_7$) could occur, depending on the sampling of actual alleles at the locus. The former, which corresponds to both identical-by-descent alleles also having the same state A_i, occurs with probability p_i^2, whereas the latter, which corresponds to distinct A_i and A_j alleles, occurs with probability $2p_ip_j$. Table 1 lists the probability of observing each of these patterns of identity-in-state conditioned upon the two individuals being related according to each of the possible modes of identity-by-descent (Figure 1), $\Pr(\mathcal{S}_i|S_j)$. Note that this table corrects a typographical error present in the classical formulation (Thompson 1975). Ritland (2000) gives a more compact notation, which may be useful computationally, that covers the rightmost three columns of Table 1; however, the compactness also makes the biological structure of the model less apparent.

Recall that the set of parameters $\Delta = (\Delta_1, \Delta_2, \ldots \Delta_9)$ correspond to the probabilities of each identity-by-descent mode and completely quantify the degree of relatedness between individuals. Following Thompson (1975), the probability of observing a particular allelic pattern, $\mathcal{S}_i$, for two individuals at a single locus, given the degree of relatedness Δ and the distribution of allele frequencies, is equal to the likelihood of Δ:

$$\begin{aligned} L(\Delta) &= \Pr(\mathcal{S}_i|\Delta) \\ &= \sum_j \Pr(\mathcal{S}_i|S_j)\Delta_j. \end{aligned} \qquad (2)$$

The likelihood of the entire sampled array of L loci is simply the product of Equation 2 across loci. Although each locus will be characterized by its own set of allele frequencies, the degree of relatedness between the two individuals (the parameter Δ in Equation 2) is constant across loci as it represents the overall relatedness of the *individuals* to each other.

Parameter space: The maximum-likelihood estimate of the set of Δ is found by searching over the parameter space until a maximum is found. In general an algebraic solution is impossible; as a result numerical methods are used. The implementation used here is based on a translation of the simplex method (Press *et al.* 1992), a hill-climbing optimization technique, into a set of C++ classes (B. G. Milligan, unpublished data). Although it is possible for such methods to identify local rather than global maxima in the likelihood function,

TABLE 1

Probability of patterns of identity-in-state $\mathcal{S}_i$ given modes of identity-by-descent S_j

Identity-in-state mode	Allelic state		Identity-by-descent mode S_j								
			S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9
$\mathcal{S}_1$	A_iA_i, A_iA_i	$\forall_i$	p_i	p_i^2	p_i^2	p_i^3	p_i^2	p_i^3	p_i^2	p_i^3	p_i^4
$\mathcal{S}_2$	A_iA_i, A_jA_j	$\forall i, \forall j \neq i$	0	p_ip_j	0	$p_ip_j^2$	0	$p_i^2p_j$	0	0	$p_i^2p_j^2$
$\mathcal{S}_3$	A_iA_i, A_iA_j	$\forall i, \forall j \neq i$	0	0	p_ip_j	$2p_i^2p_j$	0	0	0	$p_i^2p_j$	$2p_i^3p_j$
$\mathcal{S}_4$	A_iA_i, A_jA_k	$\forall i, \forall j \neq i, \forall k > j, k \neq i$	0	0	0	$2p_ip_jp_k$	0	0	0	0	$2p_i^2p_jp_k$
$\mathcal{S}_5$	A_iA_j, A_iA_i	$\forall i, \forall j \neq i$	0	0	0	0	p_ip_j	$2p_i^2p_j$	0	$p_i^2p_j$	$2p_i^3p_j$
$\mathcal{S}_6$	A_jA_k, A_iA_i	$\forall i, \forall j \neq i, \forall k > j, k \neq i$	0	0	0	0	0	$2p_ip_jp_k$	0	0	$2p_i^2p_jp_k$
$\mathcal{S}_7$	A_iA_j, A_iA_j	$\forall i, \forall j > i$	0	0	0	0	0	0	$2p_ip_j$	$p_ip_j(p_i + p_j)$	$4p_i^2p_j^2$
$\mathcal{S}_8$	A_iA_j, A_iA_k	$\forall i, \forall j \neq i, \forall k \neq i, j$	0	0	0	0	0	0	0	$p_ip_jp_k$	$4p_i^2p_jp_k$
$\mathcal{S}_9$	A_iA_j, A_kA_l	$\forall i, \forall j > i, \forall k \neq i, j,$ $\forall l > k, l \neq i, j$	0	0	0	0	0	0	0	0	$4p_ip_jp_kp_l$

Alleles with different labels (*e.g.*, A_i and A_j) are distinct, and the frequency of allele A_i is p_i.

plots of the likelihood surface and evaluation of the algorithm from distinct initial conditions suggest that multiple optima are unlikely to exist, in agreement with the analytical results of Thompson (1975).

A number of possibilities exist for defining the parameter space over which the optimization will be carried out. The complete parameter space is, of course, eight-dimensional, corresponding to the nine distinct parameters $\Delta_j = \Pr(S_j)$ constrained by the fact that they sum to unity. The immediate purpose, however, is to consider the case of a large noninbred population. In this instance only the last three parameters are nonzero. For the purposes of this analysis, the maximum-likelihood estimate is obtained by optimization within the two-dimensional parameter space defined by the parameters $(\Delta_7, \Delta_8, \Delta_9)$ constrained by their sum being unity. It is meaningless to admit solutions outside this region as they correspond to undefined values for the probability of identity-by-descent (Thompson 1975).

One of the useful features of maximum-likelihood estimators is that they can be readily adapted to a variety of situations. In some cases it may be known that individuals are either full-sibs or unrelated (or any other pair of degrees of relatedness; Mousseau *et al.* 1998). In such a case, the likelihood could be maximized within the one-dimensional parameter space representing a continuum of linear combinations of those two degrees of relatedness. Alternatively, it may be known that individuals are either full-sibs, half-sibs, or unrelated or full-sibs, parent-offspring, or unrelated (Thomas *et al.* 2002). In such cases, the optimization can be done within the parameter space defined by appropriate linear combinations of these three degrees of relatedness. Thus, the basic-likelihood estimator described here can easily be adapted to a variety of different population structures simply by appropriate choice of parameter space.

Of primary concern is the statistical behavior of the likelihood estimator, especially in contrast to existing alternatives (Queller and Goodnight 1989; Li *et al.* 1993; Ritland 1996a; Lynch and Ritland 1999; Wang 2002). Because of the reliance on numerical optimization, the statistical behavior must be determined by simulation. The basic simulation process involved generating replicate genetic data sets for a pair of individuals under conditions of known relatedness, a specified number of sampled loci, and a known distribution of allelic variation at each locus. Analytical results based on appendix a were used to verify the simulations.

Method-of-moments estimators: The performance of 6 estimators of relatedness, quantified as the coancestry coefficient θ, was investigated. The likelihood estimator calculated θ from Equation 1 as the maximum-likelihood estimate of the identity-by-descent probabilities, Δ. Five additional nonlikelihood estimators were considered as being representative of the diversity of the ones available. They represent 5 different means of using the similarity in allelic states between individuals to construct estimates of relatedness, and the ones tested by Van de Casteele *et al.* (2001) performed favorably under some conditions in their evaluation of 10 different estimators. These have all been described and compared previously (Lynch and Ritland 1999; Van de Casteele *et al.* 2001; Wang 2002); consequently, rather than repeating their formulations here the reader is referred to the specific equations presented by Ritland (1996a), Lynch and Ritland (1999), and Wang (2002).

The most commonly used estimator is one published by Queller and Goodnight (1989), of which a number of variants are possible. The variant chosen here was the symmetric one obtained by averaging $(\hat{r}_{xy} + \hat{r}_{yx})/2$ (Lynch and Ritland 1999, Equation 11) across loci. A second estimator of relatedness (Li *et al.* 1993) is based directly on the pattern of shared alleles between two individuals and was obtained by averaging $\hat{r}_{xy}$ (Lynch and Ritland 1999, Equation 8) across loci. The third estimator, a method-of-moments one based on the correlation between relatedness and genotypic similarity

(Ritland 1996a), was obtained as $\hat{\rho}$ (Ritland 1996a, Equation 5). The fourth estimator, another method-of-moments one based on the regression of similarity on relatedness (Lynch and Ritland 1999), was the weighted average of $(\hat{r}_{xy} + \hat{r}_{yx})/2$ across loci (Lynch and Ritland 1999, Equations 5–7). The final estimator was described by Wang (2002). Two forms of this were investigated. The first follows the method favored in this article and involves the following steps: calculating a set of locus-specific similarity and allele-frequency parameters, calculating a set of locus-specific weights from those locus-specific parameters, averaging the locus-specific parameters across loci using those weights, and averaging the locus-specific values of relatedness (obtained using the average parameters; Wang 2002, Equations 6 and 7 or 9–11) across loci, again using the same weights. The second method involves calculating the locus-specific values of relatedness using the locus-specific parameters (rather than the averages) and averaging across loci in the same way. These two methods are identical when allele frequencies are the same across loci; however, they may differ otherwise. Although seemingly more natural, apparently the latter approach performed less well than the former (J. Wang, personal communication). Note that some of these nonlikelihood estimators are of $r = 2\theta$ and so were transformed to θ to be comparable with the maximum-likelihood estimator.

Several of these estimators have undesirable behavior under certain conditions. For example, with two alleles the Queller and Goodnight (1989) estimator is undefined for heterozygous reference individuals, and for two equally frequent alleles the Lynch and Ritland (1999) estimator is also undefined for heterozygous reference individuals. Consequently, some loci must be discarded for multilocus estimates under these conditions, depending on the sampling of alleles at each locus. Both of these estimators are also based on arithmetic averages using each of the two individuals as a reference. At a particular locus one, both, or neither of those individuals can be heterozygous. Thus, the locus-specific value actually used in averaging across loci was the average (as described above) if both locus-specific values are defined and the defined value if only one is defined; otherwise, the locus was ignored. This approach attempts to maximize the amount of information obtained from these estimators within the constraints imposed by their mathematical definition.

Additionally, the method-of-moments estimators are generally not constrained to lie within the biologically relevant range of [0, 0.5], unlike the traditional maximum-likelihood estimator. This property enables them to remain statistically unbiased; however, individual estimates may not have meaning when interpreted as probabilities of identity-by-descent. One means of handling this is to truncate the method-of-moments estimates to lie within the proper range, that is, to replace lower values with zero and larger values with 0.5. To investigate the effect of the parameter range itself, as opposed to the type of estimator, all method-of-moments estimators were examined in both the standard and the truncated form.

Simulations: A range of sampling conditions was considered to mimic the variety of different genetic markers that might be available for estimating relatedness. Although for some organisms huge arrays of polymorphic genetic loci are available, for the vast majority of natural populations this is not the case. Thus, the range in number of loci (5–30) mimics moderate genetic samples. A great variety of types of genetic markers are available for quantifying relatedness. These range from markers that segregate few distinguishable alleles, generally including allozymes, single-nucleotide polymorphisms, or restriction/PCR fragments, to markers that segregate many distinguishable alleles, commonly microsatellites. To reflect this range in marker types, loci segregating 2, 5, 10, and 20 alleles were considered. Three different allele-frequency distributions were used for the simulations: one in which all alleles occur at equal frequency, one in which a single allele occurs with a frequency of 0.8 and the remainder are equally frequent, and one in which allele frequencies at each locus were independently drawn from the same Dirichlet distribution (Stuart and Ord 1987, Exercise 5.33, p. 209) with all parameters set to unity. The last case approximates natural situations better by allowing the distributions to vary across loci; however, the first two are useful for isolating the effects of each factor. Finally, four actual degrees of relatedness between individuals were considered: parent-offspring, full-sibs, first cousins, and unrelated individuals. This range of conditions in relatedness, numbers of loci, numbers of alleles, and types of allele-frequency distributions was chosen to make the diversity of natural situations tractable, so that the influence of each main characteristic of a genetic sample on the statistical performance of the estimators can be examined. Future studies can focus more specifically on particular types of genetic markers or population structures.

To determine the statistical behavior of each estimator under each condition, sets of 1000 replicate samples of two individuals were obtained. Each of the six estimators was used to estimate θ for each of the replicate samples. The mean and standard error of the population of estimates were calculated from these samples. The bias of each estimator was quantified as the deviation of the mean from the known parametric value of θ used to generate the data. The root mean-square error was quantified as

$$\mathrm{RMSE}(\theta) = \sqrt{\frac{1}{1000}\sum_{i=1}^{1000}(\hat{\theta}_i - \theta^*)^2}, \qquad (3)$$

where $\hat{\theta}_i$ is the ith estimate and θ^* is the parametric value used to generate the simulated data sets. In all

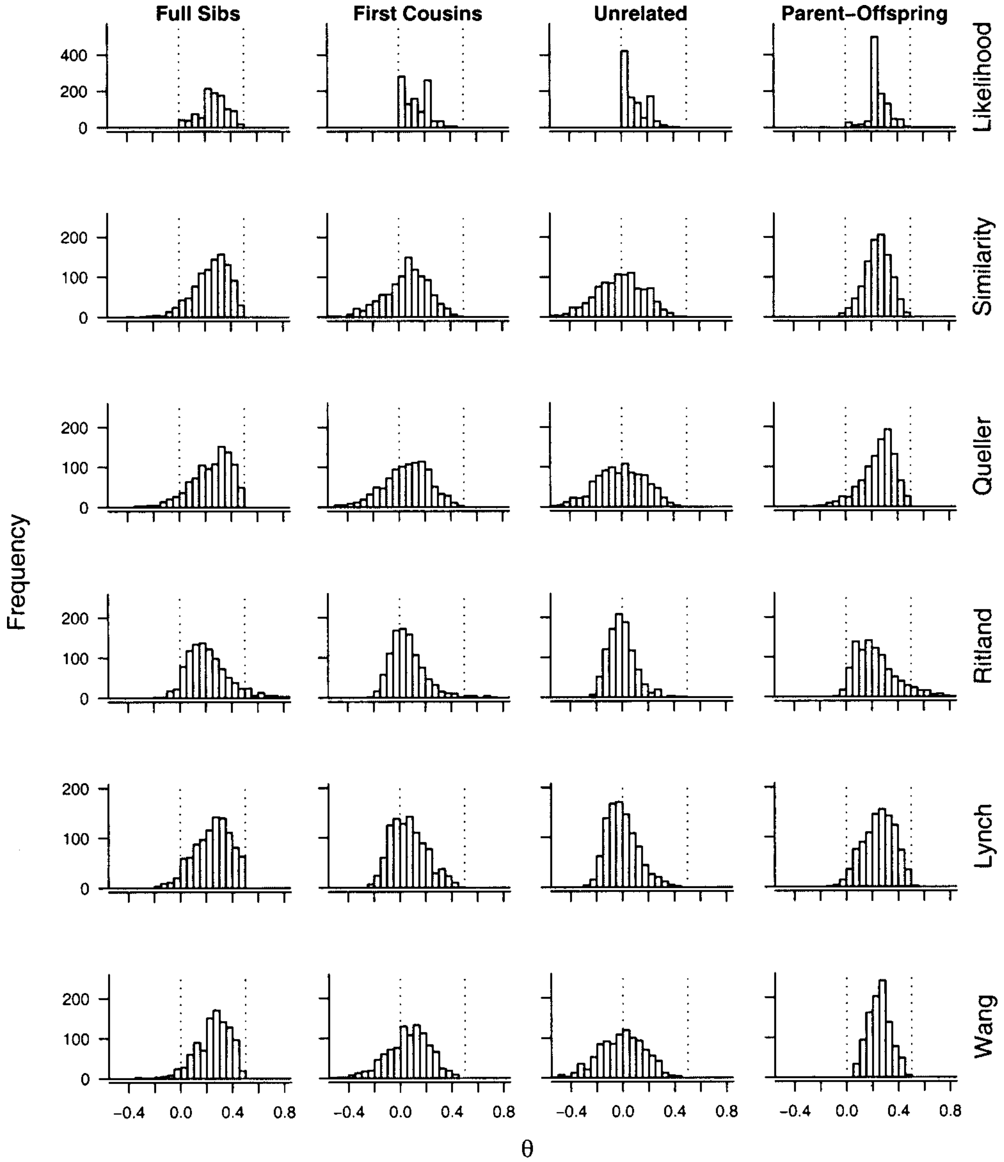

FIGURE 2.—Distribution of estimates of θ for each estimator under different conditions of true relatedness. Each distribution is based on a sample of 1000 estimates taken from five loci, each segregating for five alleles. The allele-frequency distributions for each locus are drawn from a Dirichlet distribution. The WANG (2002) estimator is represented by the published version.

cases it was assumed that the allele-frequency distribution was known without error. Consequently, the focus is on the sampling properties of the relatedness estimators themselves.

RESULTS

As with any estimator of genetic relatedness, the quality of the estimate depends on the amount of available genetic information. Typically, both the number of loci for which genetic information is available and the number of alleles segregating at those loci have strong influences on the standard error of the estimate of relatedness. Additionally, different estimators of relatedness often respond differently to the amount of genetic information available.

Figure 2 illustrates the general level of variation yielded by each of the estimators. It is evident that the likelihood estimator described here has lower standard error than any of the others under all conditions; however, the other five are rather similar overall. One feature of the other estimators is their propensity to yield estimates of θ that lie outside the biologically meaningful range of [0, 0.5]. Under some conditions approximately half of the estimates are negative, for example. Because the focus of interest for these measures is on a specific pair of individuals, it is difficult to interpret the meaning of estimates that suggest, for example, that two individuals are less related than unrelated. However, because the likelihood estimator is constrained to always produce estimates within the biologically meaningful range, some bias is introduced near the boundary. For example, for unrelated individuals $\theta = 0$, yet the likelihood estimator evidently commonly generates values that overestimate that. Thus, while exhibiting less variation, the likelihood estimator is more biased under some conditions. Clearly, the truncated estimators will also exhibit less variation and more bias than the untruncated ones for the same reason.

An additional feature that is evident from Figure 2 is

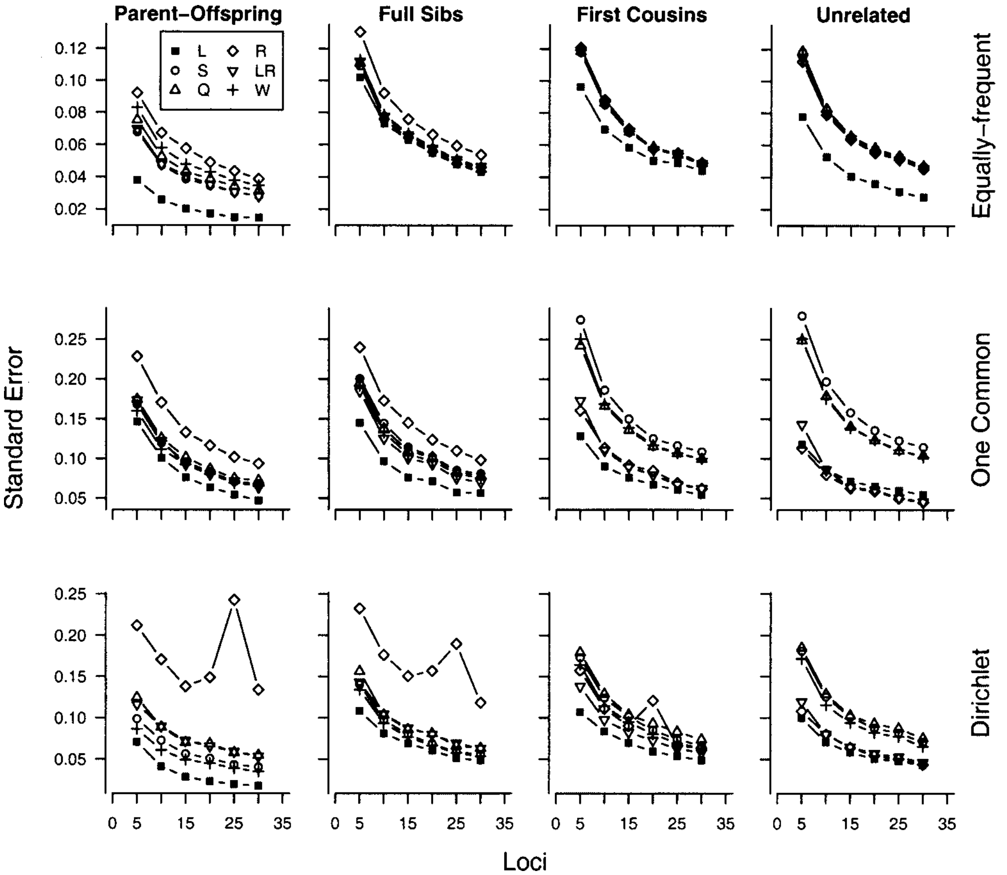

FIGURE 3.—Comparison of standard error among estimators. These curves are based on samples of 1000 pairs of individuals, each related as shown, assayed for different numbers of loci, each of which segregates for five alleles. The top corresponds to five equally frequent alleles, the middle corresponds to one common allele and four rare ones, and the bottom corresponds to allele frequency distributions drawn from a Dirichlet distribution.

that many of the estimators are skewed. This is particularly the case for the QUELLER and GOODNIGHT (1989) and especially the RITLAND (1996a) estimators, which interestingly are skewed in opposite directions. This skew may have a significant impact on the use of these estimators, because even though they are essentially unbiased in expectation, the modal estimate does not equal the expectation; the most probable outcome in any particular case will be an incorrect estimate.

Standard error: Figure 3 quantifies the standard error of each estimator of θ as a function of the amount of genetic information available. The variation for all estimators declines with the number of loci sampled. Generally, the standard error of the likelihood estimator is lower than that of any of the others. Interestingly, the estimators proposed by RITLAND (1996a) and LYNCH and RITLAND (1999) approach the likelihood estimator under conditions of both low degree of relatedness and very restricted genetic information. Undoubtedly, this is because the weights used across loci for these estimators assume no relatedness. However, their performance is not consistent even for unrelated individuals; when more genetic information is available in the form of more uniform allele-frequency distributions, their performance is consistent with the other nonlikelihood estimators. The RITLAND (1996a) estimator even performs distinctly worse than any other under some conditions considered. This anomalous behavior is because it involves a ratio of the allelic similarity (a number from the set {0, 0.25, 0.5, 1}) and the allele frequency. If by chance a rare allele is shared between individuals, this ratio can be substantial, greatly inflating the variance of the estimator and causing the extreme skewness observed in Figure 2. Thus, the utility of this estimator is highly dependent on the unknown quantity being estimated, relatedness, and the details of the genetic sample. To a lesser degree, the same is true of the LYNCH and RITLAND (1999) estimator. In contrast, the likelihood estimator maintains a low standard error across the full range of conditions. As such, it may be preferable overall, because in general no information is available to enable one to choose among the estimators on the basis of their performance under the special conditions of actual relatedness applying to a particular pair of individuals.

In these simulations there is no indication that the two different versions of the WANG (2002) estimators differ substantially. Both are essentially unbiased under all conditions and only for parent-offspring pairs were the differences between them >1%. Nor is there a strong indication that they represent substantial improvements over the other estimators. In no case is either better than the likelihood estimator. Perhaps these results are a consequence of the slightly different statistics used to evaluate the estimators. Whereas LYNCH and RITLAND (1999) and WANG (2002) use the single-locus variance, quantified as a mean for 10-locus samples, to evaluate the estimators, Figure 3 evaluates them on the

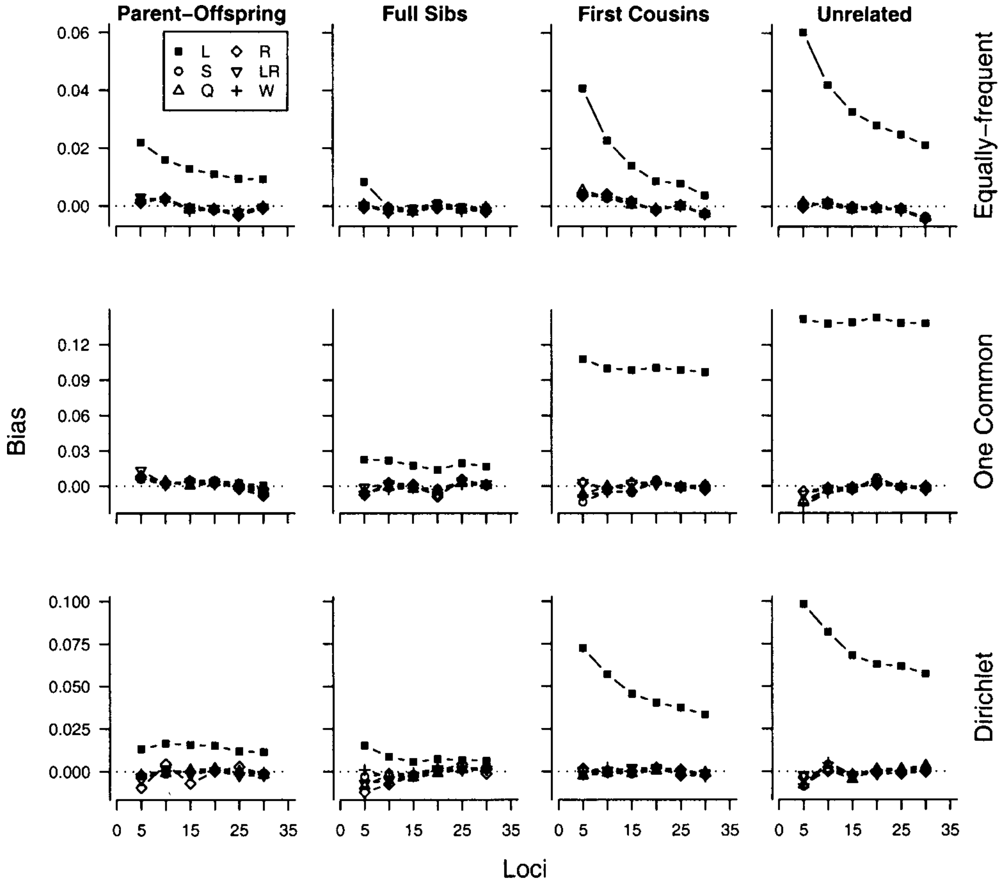

FIGURE 4.—Comparison of bias among estimators. These curves are based on the same samples used to construct Figure 3.

basis of the actual variance for many different sample sizes. This approach is adopted for two important reasons. First, it more closely approximates the information needed directly for experimental design purposes, when the fundamental decision is often based on the reduction of variance as a function of genetic sample size. Second, the relative performance of the different estimators is a function of the number of loci sampled; basing an evaluation on a single sample size may be misleading for extrapolations to another.

The actual degree of relatedness between individuals has relatively little effect on the standard error of the likelihood estimator of θ. For example, the standard errors of 30-locus likelihood estimates for full-sibs and first cousins differ by <4%, despite these representing quite different degrees of relatedness. The standard error is lower for unrelated individuals because of the constraint that estimates must be within the biologically realistic range. This independence of actual relatedness for standard error of θ is broadly consistent across all the estimators when the allele-frequency distribution is favorable; it is somewhat less so when the allele-frequency distribution is dominated by a single allele segregating at high frequency or when variation among loci in allele-frequency distributions exists.

Bias: The second main statistical feature of each estimator, bias, is illustrated in Figure 4. Two features are immediately apparent. All of the nonlikelihood estimators are essentially unbiased under all conditions; in contrast, the likelihood estimator is biased under some conditions. As mentioned before, this is a consequence of the fact that the likelihood estimator is constrained within the biologically meaningful range of [0, 0.5].

Unlike for standard error, the actual degree of relatedness does influence the bias of the likelihood estimator. For example, the likelihood estimator for parent-offspring and full-sib relationships yields estimates that are quite close to the true value of θ; in fact, its bias is either zero or close enough as to be biologically insignificant. However, when actual relatedness is close to the boundary, which is the case for both first cousins and unrelated individuals, the bias is much larger. This is true for any of the allele-frequency distributions. When the allele-frequency distribution is favorable, increasing numbers of loci can substantially reduce the bias of the likelihood estimator, potentially to the point of being biologically insignificant even for unrelated individuals. In contrast, when the allele-frequency distribution is dominated by a single allele, increasing numbers of loci have little effect on bias for reasonable numbers of loci.

Root mean-square error: The third main statistical feature of each estimator, root mean-square error, is illustrated in Figure 5. This measure is a reflection of the mean deviation of the distribution of estimates from the parametric value of θ used in the simulation. As

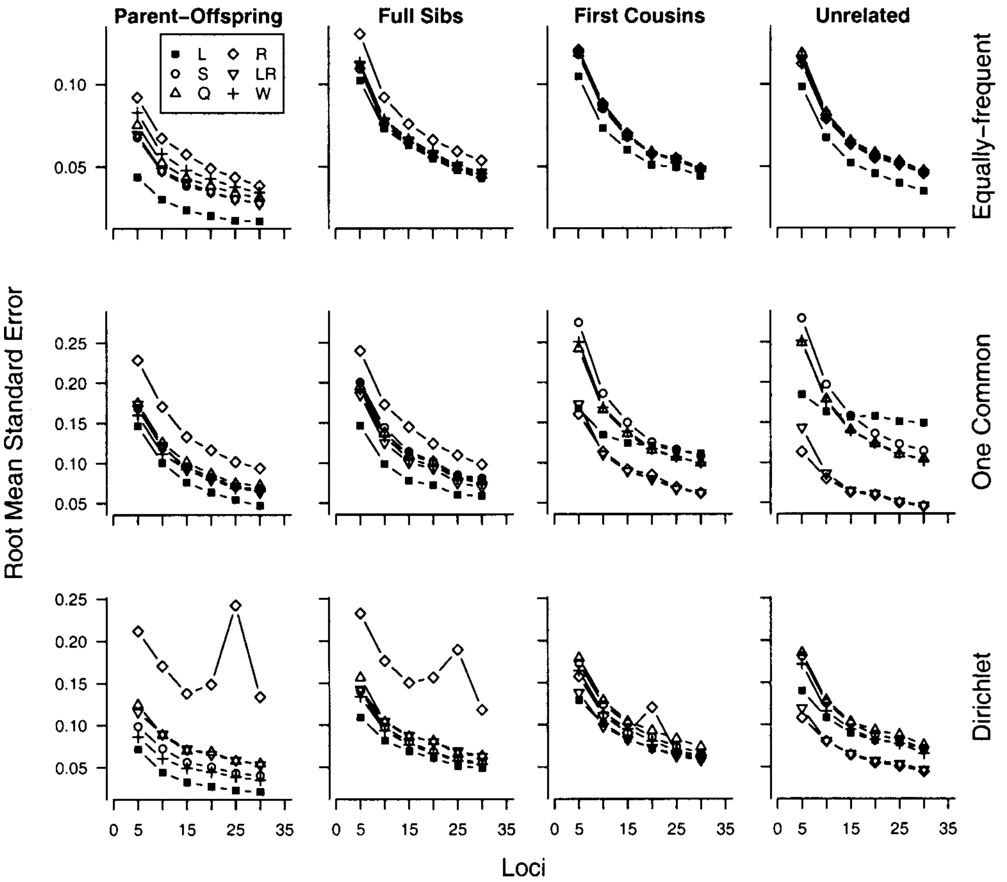

FIGURE 5.—Comparison of root mean-square error among estimators. These curves are based on the same samples used to construct Figure 3.

such it integrates both the standard error and the bias of the estimators. Largely these curves follow the corresponding ones for standard error, an indication that under most conditions all estimators are essentially unbiased. Only in situations both lacking in useful genetic information (*e.g.*, a single predominant allele at each locus) and with true relationships near the boundary of the parameter space (*e.g.*, especially unrelated individuals) does the likelihood estimator perform notably worse than the others with regard to root mean-square error. In all other cases considered, the likelihood estimator performs better than any alternative with regard to this integrated measure of performance.

Truncated estimators: The performance of the truncated method-of-moments estimators relative to the likelihood estimator is largely anticipated from Figure 2. In the cases of full-sib and parent-offspring relatedness relatively few estimates lie beyond the meaningful range of [0, 0.5], so truncation has little effect; in contrast, for first cousins and unrelated individuals, substantial numbers of estimates do so and truncation has a large effect. For example, although the standard error is essentially unchanged between truncated and nontruncated estimators for the former two, it is somewhat reduced for first cousins and substantially reduced for unrelated individuals. For first cousins the truncated method-of-moments estimators exhibit no standard error lower than that of the maximum-likelihood estimator, and in most cases still exceed it. The reduction in standard error is great enough for unrelated individuals that, for the Dirichlet distribution of allele frequencies and when one allele predominates the RITLAND (1996a) and LYNCH and RITLAND (1999) estimators, both exhibit a lower standard error than that of the maximum-likelihood estimator. The others, and all truncated estimators when alleles are equally frequent, exhibit standard errors similar to the maximum-likelihood estimator. Interestingly, the relative ranking of the method-of-moments estimators is quite similar whether they are truncated or not.

The bias of the truncated method-of-moments estimators increases substantially for the two less-related cases. Although the relative increase in bias is quite large (>10- to 100-fold in some cases), the bias is somewhat less than that exhibited by the maximum-likelihood estimator. Under these conditions the truncated RITLAND (1996a) and LYNCH and RITLAND (1999) estimators exhibited the lowest bias. However, the truncated RITLAND (1996a) estimator is notably more biased under the two closer degrees of true relatedness, even exceeding the bias of the maximum-likelihood estimator under some conditions. This is another manifestation of the sensitivity of this estimator to the conditions under study, at least some of which will be unknown in any realistic situation.

Segregating alleles: The previous discussion illustrated the performance of alternative estimators under several different sampling conditions. These results are

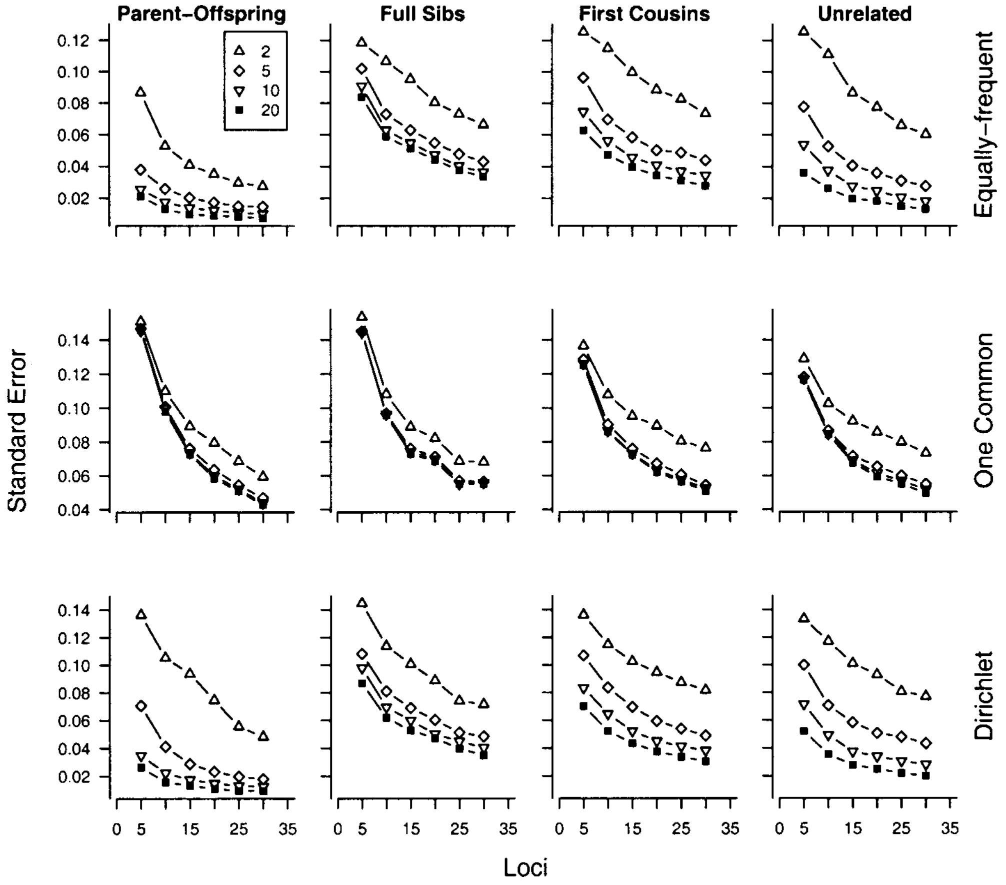

Figure 6.—Dependence of the standard error of the likelihood estimator on both the number of loci sampled and the number of alleles segregating at each. Each point is based on 1000 replicate samples. Symbols indicate different numbers of alleles segregating.

quite typical. However, the difference between the likelihood estimator and the others declines as the number of alleles segregating at each locus increases. That is, the nonlikelihood estimators improve in performance and approach the likelihood estimator as the amount of genetic information increases.

The variation of the likelihood estimator itself also changes in response to increasing numbers of alleles segregating at each locus (Figure 6). In all cases, more segregating alleles reduce the standard error of the likelihood estimate of θ. This is especially true when alleles are equally frequent. When one allele predominates, even a few additional alleles substantially reduce the standard error, which is not further reduced by many additional alleles. Thus, for a wide range of conditions a large reduction in standard error relative to biallelic samples is possible by sampling loci with even a few alleles; additional reductions are possible only if many alleles segregate at intermediate frequencies.

In contrast, the bias of the likelihood estimator is relatively unaffected by the number of alleles segregating at each locus (Figure 7). In fact, when the allele-frequency distribution is dominated by a single allele, the number of additional alleles segregating (and, as noted above, the number of loci) has essentially no influence on the bias over the range of loci considered (although the estimator is asymptotically unbiased). These sampling conditions are basically uninformative, so large samples are unhelpful. However, if the allele-frequency distribution is more favorable, the number of alleles segregating at each locus does have an influence on the bias of the likelihood estimator. As with the standard error, substantial reductions in bias are possible under all conditions of actual relatedness when even a few alleles are segregating. Even for conditions exhibiting the largest bias (*e.g.*, unrelated individuals), the degree of bias can be reduced to biologically insignificant levels for realistic samples.

DISCUSSION

Despite its basic importance for understanding the biology of natural populations, estimation of relatedness between individuals remains a difficult challenge. A diversity of estimators (Thompson 1975; Queller and Goodnight 1989; Li *et al.* 1993; Ritland 1996a; Lynch and Ritland 1999; Wang 2002) have been developed to use the information contained within samples of molecular markers to estimate relatedness. With one exception (Thompson 1975), none of the methods proposed to date are traditional maximum-likelihood estimators. Here we investigate its statistical properties in comparison with five of the commonly used alternatives.

The prominent feature of the likelihood estimator is

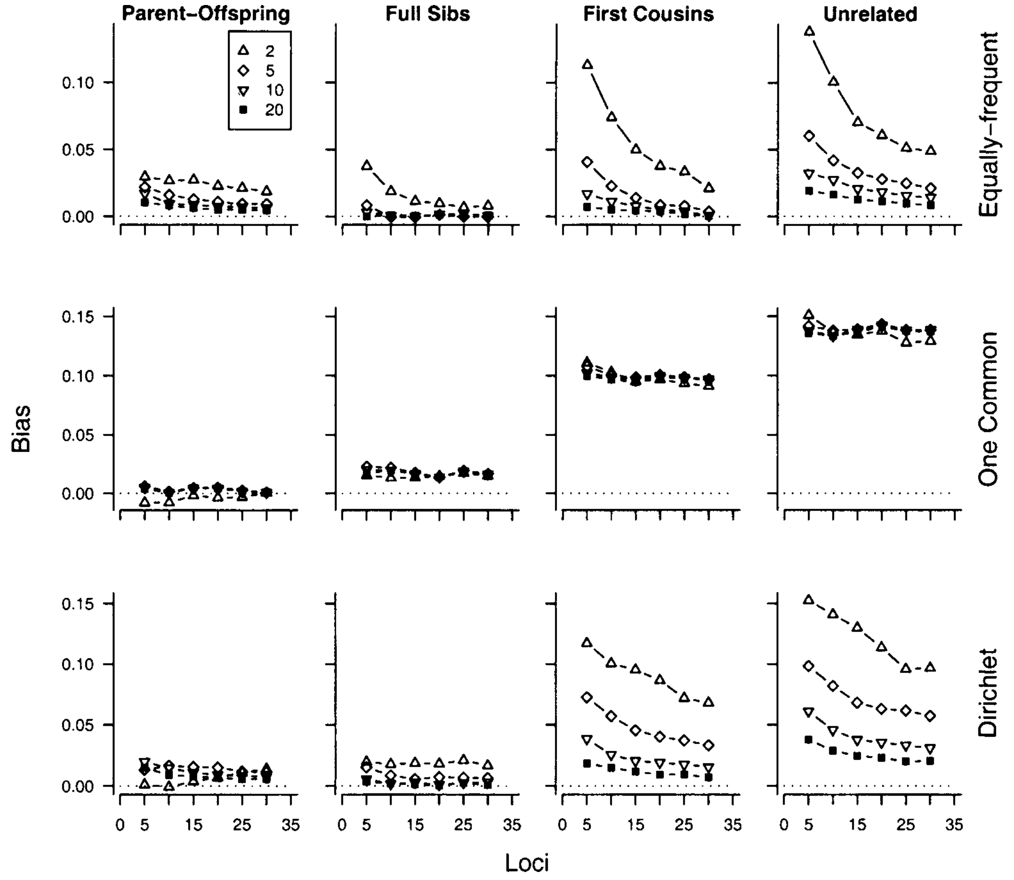

Figure 7.—Dependence of the bias of the likelihood estimator on both the number of loci sampled and the number of alleles segregating at each. These are based on the same samples used to obtain Figure 6. Symbols indicate different numbers of alleles segregating.

that it exhibits a lower standard error than that of any of the others for a wide range of reasonable sampling conditions. This conclusion appears to be independent of the number of loci sampled, the number of alleles segregating at each locus, or the frequency distribution of the segregating alleles. When many loci are sampled, the other estimators approach the performance of the likelihood estimator. This is especially true for the Ritland (1996a) and Lynch and Ritland (1999) estimators when individuals are only distantly related and the allele-frequency distribution is highly skewed. However, especially the Ritland (1996a) estimator is much less consistent across different sampling conditions and exhibits a strong dependency on the actual degree of relatedness, the unknowable quantity that is being estimated. Furthermore, the standard error of the likelihood estimator is largely independent of the actual degree of relatedness between individuals. Thus, from this standpoint the likelihood estimator exhibits preferable statistical behavior compared with any of the alternatives, because it is the only one that consistently maintains a low standard error across all conditions.

These conclusions differ dramatically from those obtained earlier for different maximum-likelihood estimators of relatedness (Ritland 1996a; Lynch and Ritland 1999). Their estimators performed so poorly as to be immediately discarded as useless in practice. This difference arises either from admitting solutions that cannot be directly interpreted biologically as probabilities of identity-by-descent or from the nature of the likelihood function (see appendix b). In contrast, the likelihood estimator investigated here is consistent with the traditional literature on likelihood estimation of relatedness (Thompson 1975) and admits only solutions that are fully interpretable biologically. As a result, it performs much better than previously suggested for maximum-likelihood estimators.

The other feature of the likelihood estimator is that it, unlike the others, is biased under some conditions. The degree of bias is dependent on the actual degree of relatedness between individuals and the nature of the genetic information. If the actual degree of relatedness is near a boundary, such as for first cousins or unrelated individuals, the bias is more severe than if the actual degree of relatedness is within the interior of the parameter space, such as for full-sibs. However, the degree of bias can be greatly reduced by sampling loci that segregate for more alleles. Additional segregating alleles are not helpful if their frequency distribution is highly skewed. Samples of 20–30 microsatellite loci, which often segregate for 20 or more alleles and exhibit high heterozygosity, could yield estimates of θ that are quite unbiased, even for unrelated individuals. However, even markers segregating for only a few alleles can also dramatically reduce the bias compared with biallelic markers. Thus, even though the likelihood estimator is more biased than any of the others, suitable genetic sampling can greatly reduce this problem. The amount of genetic information required to reduce the bias of the likelihood estimator to insignificant levels is within the range of feasible sampling efforts.

These two quantities, standard error and bias, are integrated by the measure of root mean-square error. As with the standard error, the likelihood estimator maintains a low root mean-square error under almost all conditions. The exceptions are cases involving little useful genetic information and true relationships that are near the boundary of the parameter space. However, the fact that a low root mean-square error is maintained by the likelihood estimator, despite the inherent potential for enhanced bias, indicates that in fact the bias is of little biological consequence. The other estimators more than make up for generally lower bias through their greater standard error. Additionally, the skewness identified for several other estimators may lead to further problems in practice.

Often the primary interest lies in the estimate of relatedness itself. Such would be the case, for example, in ascertaining family membership or determining the spatial structure of relatedness for sessile organisms. In such cases it is critical that each estimate of relatedness yield a biologically meaningful value. Such is the case for the likelihood estimator, which is constrained to yield estimates of θ within the biologically meaningful range of [0, 0.5]. The method-of-moments estimators may also be truncated to lie within the same range. Under conditions of low relatedness, this reduces their standard error and increases their bias. In many cases, however, the maximum-likelihood estimator still exhibits lower standard error or root mean-square error than that of the truncated ones. Thus, the general performance characteristics of this estimator are not strictly the result of constraining the parameter space to include only biologically meaningful estimates.

The primary interest in estimating relatedness may alternatively lie in using the estimates in a subsequent analysis. Such would be the case, for example, in estimating heritability or additive genetic variance from relatedness estimates and phenotypic observations (Ritland 1996b, 2000; Ritland and Ritland 1996; Mousseau *et al.* 1998; Thomas *et al.* 2000, 2002). In these cases, too, it may be problematic to allow relatedness estimates that lie beyond the biologically meaningful range. However, the amount of variation in estimates of θ can also be especially important, because of the propagation of error in the estimate of θ to variation in derived estimates of additive genetic variance, V_A.

Depending on the sampling conditions, standard errors for the nonlikelihood estimators are between 2 and 250% larger than the standard error for the likelihood estimator. That discrepancy is substantially improved only by truncating the nonlikelihood estimators when individuals are unrelated; unfortunately, in practice it will be unknown whether the specific pair of interest is unrelated or not.

This difference in standard error between estimators is likely to be significant. For example, in the study by Ritland and Ritland (1996) involving a sample of eight loci in Mimulus, the actual variance of relatedness was estimated to be only 0.04; almost all the observed variance in estimates of r was due to sampling error. Indeed, the general lack of application of marker-based estimates of relatedness to subsequent estimation of quantitative genetics parameters such as heritability or additive genetic variance may be attributable to the relatively poor performance of the available estimators. An estimator of relatedness, like the likelihood one discussed here, that exhibits both lower standard error and lower root mean-square error across many sampling conditions would improve the discriminatory power of such analyses.

Overall, the likelihood estimator discussed in this article offers several advantages compared with existing, commonly used estimators. First, except for some truncated estimators applied to unrelated individuals, it uniformly exhibits lower variation, even under conditions of relatively abundant genetic information. For example, even when 30 loci, each segregating for 20 equally frequent alleles, are sampled, the standard error is 10% (and >250% under some conditions) greater for some estimators than for the likelihood estimator. Second, all likelihood estimates are constrained to lie within the biologically meaningful range. Thus, the biological interpretation of individual estimates is quite straightforward. Although apparently not general practice, this constraint can be obtained by truncating the nonlikelihood estimates. While each estimate is now interpretable biologically, the statistical behavior of the truncated estimators is not generally improved over the maximum-likelihood estimator and in some cases is worse. Finally, the likelihood estimator naturally accommodates different genetic sampling schemes. For example, the relative weighting of data from microsatellite loci segregrating for many alleles in contrast to data from single-nucleotide polymorphisms segregating for only two alleles is accomplished directly by the likelihood function. Consequently, all available data can be used to ascertain the degree of relatedness between two individuals.

The main drawback associated with the likelihood estimator is that it can be biased under some conditions, especially if the true degree of relatedness is near the boundary of the parameter space and little genetic information is available. The same is also true of the truncated estimators, some of which are more biased than the likelihood function even when the true relatedness is not near the boundary. Even though biased, however, the maximum-likelihood estimator exhibits a lower root mean-square error than do alternative estimators under many conditions. Thus, the bias is quite minor from a biological perspective. Furthermore, the extent of the bias can be greatly reduced by suitable genetic sampling.

If markers with many alleles segregating at intermediate frequencies are used, the bias can be reduced considerably. Often markers with only a few alleles segregating at intermediate frequencies approach the performance of highly polymorphic ones. Given the relative ease with which genetic information can be obtained, bias is not likely to be a major drawback in practice. However, a preliminary study aimed at defining the allele-frequency distributions prior to selecting loci for more intense sampling can dramatically reduce the genetic information required to obtain estimates of relatedness.

Although the maximum-likelihood estimator of relatedness performs extremely well overall, there are conditions under which others perform better according to specific metrics. None perform uniformly better under all conditions according to all metrics. Whether the most relevant performance metric is the standard error, the bias, the root mean-square error, or some other one likely depends in detail on the ultimate use of the relatedness estimates. For specific applications under specific genetic conditions it may be possible to identify one estimator that is optimal. However, for its wide range of applicability and excellent performance across almost all conditions it is difficult to improve on the maximum-likelihood estimator.

Thanks are due to Kelly Gallagher, Kermit Ritland, Elizabeth Thompson, and two anonymous reviewers for their helpful comments. Thanks are also due to Jinliang Wang for his comments and for clarifying the details of the weighting scheme used in his estimator. Finally, Sara Knott and Peter Visscher provided many stimulating discussions and insightful comments that greatly improved the manuscript. This work was supported by National Science Foundation grant DEB-9904026.

LITERATURE CITED

Boehnke, M., and N. J. Cox, 1997 Accurate inference of relationships in sib-pair linkage studies. Am. J. Hum. Genet. **61:** 423–429.

Broman, K. W., and J. L. Weber, 1998 Estimation of pairwise relationships in the presence of genotyping error. Am. J. Hum. Genet. **63:** 1563–1564.

Crow, J., and M. Kimura, 1970 *An Introduction to Population Genetics.* Burgess, Minneapolis.

Falconer, D. S., 1981 *Introduction to Quantitative Genetics*, Ed. 2. Longman, London.

Hamilton, W. D., 1964a The genetical evolution of social behaviour. I. J. Theor. Biol. **7:** 1–16.

Hamilton, W. D., 1964b The genetical evolution of social behaviour. II. J. Theor. Biol. **7:** 17–52.

Jacquard, A., 1974 *The Genetic Structure of Populations.* Springer-Verlag, New York.

Kendall, M. G., A. Stuart and J. K. Ord, 1979 *Advanced Theory of Statistics, Vol. 2: Inference and Relationship,* Ed. 4. Macmillan, New York.

Li, C. C., D. E. Weeks and A. Chakravarti, 1993 Similarity of DNA fingerprints due to chance and relatedness. Hum. Hered. **43:** 45–52.

Lynch, M., and K. Ritland, 1999 Estimation of pairwise relatedness with molecular markers. Genetics **152:** 1753–1766.

Lynch, M., and B. Walsh, 1998 *Genetics and Analysis of Quantitative Traits.* Sinauer, Sunderland, MA.

Mousseau, T. A., K. Ritland and D. D. Heath, 1998 A novel method for estimating heritability using molecular markers. Heredity **80:** 218–224.

Painter, I., 1997 Sibship reconstruction without parental information. J. Agric. Biol. Environ. Stat. **2:** 212–229.

Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, 1992 *Numerical Recipes in C: The Art of Scientific Computing*, Ed. 2. Cambridge University Press, Cambridge, UK.

Queller, D. C., and K. F. Goodnight, 1989 Estimating relatedness using genetic markers. Evolution **43:** 258–275.

Ritland, K., 1996a Estimators for pairwise relatedness and individual inbreeding coefficients. Genet. Res. **67:** 175–185.

Ritland, K., 1996b A marker-based method for inferences about quantitative inheritance in natural populations. Evolution **50:** 1062–1073.

Ritland, K., 2000 Marker-inferred relatedness as a tool for detecting heritability in nature. Mol. Ecol. **9:** 1195–1204.

Ritland, K., and C. Ritland, 1996 Inferences about quantitative inheritance based on natural population structure in the yellow monkeyflower, *Mimulus guttatus.* Evolution **50:** 1074–1082.

Sieberts, S. K., E. M. Wusman and E. A. Thompson, 2002 Relationship inference from trios of individuals, in the presence of typing error. Am. J. Hum. Genet. **70:** 170–180.

Stuart, A., and J. K. Ord, 1987 *Kendall's Advanced Theory of Statistics, Vol. 1: Distribution Theory,* Ed. 5. Oxford University Press, New York.

Thomas, S. C., J. M. Pemberton and W. G. Hill, 2000 Estimating variance components in natural populations using inferred relationships. Heredity **84:** 427–436.

Thomas, S. C., D. W. Coltman and J. M. Pemberton, 2002 The use of marker-based relationship information to estimate the heritability of body weight in a natural population: a cautionary tale. J. Evol. Biol. **15:** 92–99.

Thompson, E. A., 1975 The estimation of pairwise relationships. Ann. Hum. Genet. **39:** 173–188.

Thompson, E. A., 1986 *Pedigree Analysis in Human Genetics.* Johns Hopkins University Press, Baltimore.

Van de Casteele, T., P. Galbusera and E. Matthysen, 2001 A comparison of microsatellite-based pairwise relatedness estimators. Mol. Ecol. **10:** 1539–1549.

Wang, J., 2002 An estimator for pairwise relatedness using molecular markers. Genetics **160:** 1203–1215.

Communicating editor: J. B. Walsh

APPENDIX A

Although the likelihood function is in general complex and does not admit full analysis, some insight can be obtained for special cases and those results can be used to test the simulations. The first special case to consider corresponds to a single locus segregating for n equally frequent alleles. Table A1 gives a summary of this situation. Each of the allelic identity-in-state patterns (see Table 1) $\mathcal{S}_1$–$\mathcal{S}_9$ recurs a number of times with different alleles; for example, the pattern involving all four alleles identical-in-state ($\mathcal{S}_1$) will occur n different times, corresponding to each of the n different alleles that are possible. Given that the true relationship between the individuals is known, the probability of each of the identity-in-state patterns can be calculated. These are given for each of the fundamental modes of relationship in a noninbred population ($\Delta_7 = 1$, $\Delta_8 = 1$, and $\Delta_9 = 1$); any other is simply a linear combination of these.

For the case of a single locus segregating for n equally frequent alleles, the maximum of the likelihood func-

TABLE A1

Single-locus cases segregating *n* equally frequent alleles

		Probability of identity-in-state			Maximum-likelihood estimates			
		True relatedness			Segregating alleles (n)			
Identity-in-state mode	No. of allelic combinations	$\Delta_7 = 1$	$\Delta_8 = 1$	$\Delta_9 = 1$	2	3	4	≥5
$\mathcal{S}_1$	n	$\frac{1}{n}$	$\frac{1}{n^2}$	$\frac{1}{n^3}$	$\hat{\Delta}_7 = 1$	$\hat{\Delta}_7 = 1$	$\hat{\Delta}_7 = 1$	$\hat{\Delta}_7 = 1$
$\mathcal{S}_2$	$n(n-1)$	0	0	$\frac{n-1}{n^3}$	$\hat{\Delta}_9 = 1$	$\hat{\Delta}_9 = 1$	$\hat{\Delta}_9 = 1$	$\hat{\Delta}_9 = 1$
$\mathcal{S}_3$	$n(n-1)$	0	$\frac{n-1}{n^2}$	$\frac{2(n-1)}{n^3}$	$\{\Delta_8, \Delta_9\}$	$\hat{\Delta}_8 = 1$	$\hat{\Delta}_8 = 1$	$\hat{\Delta}_8 = 1$
$\mathcal{S}_4$	$\frac{1}{2}n(n-1)(n-2)$	0	0	$\frac{(n-1)(n-2)}{n^3}$		$\hat{\Delta}_9 = 1$	$\hat{\Delta}_9 = 1$	$\hat{\Delta}_9 = 1$
$\mathcal{S}_5$	$n(n-1)$	0	$\frac{n-1}{n^2}$	$\frac{2(n-1)}{n^3}$	$\{\Delta_8, \Delta_9\}$	$\hat{\Delta}_8 = 1$	$\hat{\Delta}_8 = 1$	$\hat{\Delta}_8 = 1$
$\mathcal{S}_6$	$\frac{1}{2}n(n-1)(n-2)$	0	0	$\frac{(n-1)(n-2)}{n^3}$		$\hat{\Delta}_9 = 1$	$\hat{\Delta}_9 = 1$	$\hat{\Delta}_9 = 1$
$\mathcal{S}_7$	$\frac{1}{2}n(n-1)$	$1 - \frac{1}{n}$	$\frac{n-1}{n^2}$	$\frac{2(n-1)}{n^3}$	$\hat{\Delta}_7 = 1$	$\hat{\Delta}_7 = 1$	$\hat{\Delta}_7 = 1$	$\hat{\Delta}_7 = 1$
$\mathcal{S}_8$	$n(n-1)(n-2)$	0	$\frac{(n-1)(n-2)}{n^2}$	$\frac{4(n-1)(n-2)}{n^3}$		$\hat{\Delta}_9 = 1$	$\{\Delta_8, \Delta_9\}$	$\hat{\Delta}_8 = 1$
$\mathcal{S}_9$	$\frac{1}{4}n(n-1)(n-2)(n-3)$	0	0	$\frac{(n-1)(n-2)(n-3)}{n^3}$			$\hat{\Delta}_9 = 1$	$\hat{\Delta}_9 = 1$

The number of different combinations for each identity-in-state mode are given, together with the total probability of observing all such patterns given true relatedness corresponding to monozygotic twins ($\Delta_7 = 1$), parent-offspring ($\Delta_8 = 1$), and unrelated individuals ($\Delta_9 = 1$). The final columns list the maximum-likelihood estimates obtained for each possible identity-in-state mode given observations on a single locus and n equally frequent alleles segregating in the population. Cases designated by $\{\Delta_8, \Delta_9\}$ correspond to situations in which the estimate is indeterminant and any linear combination of the two maximizes the likelihood.

tion can be solved analytically. In some cases (*e.g.*, $\mathcal{S}_3$, $\mathcal{S}_5$, and $\mathcal{S}_8$) the maximum depends on the number of alleles segregating. Note also that when only a few alleles are segregating not all the identity-in-state modes are possible to observe.

Even from Table A1 it is possible to understand the behavior of the maximum-likelihood estimator under more general conditions. For example, if an infinite number of alleles segregate in a noninbred population the identity-in-state pattern $\mathcal{S}_i$ ($i = \{7, 8, 9\}$) will occur only when individuals are related by identity-by-descent mode S_i. The corresponding maximum-likelihood estimate will also be $\hat{\Delta}_i = 1$. More loci will reinforce the correct estimate; in cases such as full-sibs or first cousins involving a linear combination of the three fundamental relationship modes, more loci will yield an estimate of the correct proportion of loci corresponding to each of the fundamental modes. This may provide intuitive justification for the asymptotically unbiased nature of the maximum-likelihood estimator (Kendall *et al.* 1979).

The effect of fewer alleles segregating is also evident from Table A1. In this case identity-in-state patterns $\mathcal{S}_1$–$\mathcal{S}_6$ will be observed simply because identical alleles are resampled from the finite pool. The effect of this depends on the number of alleles segregating. If only two alleles are segregating, only five of the nine possible patterns are observable. Two of these yield estimates of $\hat{\Delta}_7 = 1$, one yields an estimate of $\hat{\Delta}_9 = 1$, and the remainder yield indeterminate estimates. Given that some of these patterns can occur under any degree of relationship, it is clear that the maximum-likelihood estimate may be misleading under such situations. However, this is entirely due to the fact that the information available for the inference is itself misleading, something that no estimator can alter.

The rapid improvement in performance of the maximum-likelihood estimator with increasing number of segregating alleles is also understandable from Table A1. With only two alleles, the set of observable identity-in-state patterns is quite constrained and the maximum-likelihood estimates are less concordant with the actual mode of identity-by-descent. Even one additional allele greatly improves the concordance.

A second special case that is amenable to analysis corresponds to a parent-offspring pair assayed at L loci, each of which segregates for n equally frequent alleles. In this case the true relationship is $S_8 = 1$ and the likelihood function is given by

$$
\begin{aligned}
L(\Delta_7, \Delta_8, \Delta_9) &= \left(\frac{1}{n^2}\Delta_7 + \frac{1}{n^3}\Delta_8\right)^{l_1}\left(\frac{2}{n^2}\Delta_7 + \frac{2}{n^3}\Delta_8\right)^{l_7}\left(\frac{1}{n^3}\Delta_8\right)^{l_3+l_5+l_8} \\
&= c(n\Delta_7 + \Delta_8)^{L_1}(\Delta_8)^{L_2}, \qquad \text{(A1)}
\end{aligned}
$$

where l_i is the number of loci exhibiting identity in state pattern $\mathcal{S}_i$, $L_1 = l_1 + l_7$ and $L_2 = l_3 + l_5 + l_8$, and c is a constant of proportionality independent of Δ_i. Explicit maximization of Equation A1 yields the estimate of relationship

$$
\hat{\Delta}_7 = \begin{cases} \dfrac{nL_1 - L}{(n-1)L} & \text{if } nL_1 > L \\ 0 & \text{otherwise,} \end{cases} \qquad \text{(A2)}
$$

where $L = L_1 + L_2$ is the total number of loci sampled. $\hat{\Delta}_8 = 1 - \hat{\Delta}_7$ and $\hat{\Delta}_9 = 0$. These estimates, together with the probabilities of observation given in Table A1 and the binomial distribution, can be used to derive the mean and variance of θ.

APPENDIX B

The method-of-moment estimator used by Lynch and Ritland (1999) is apparently from the same probability model as described above, based on both its derivation and its performance in simulations. However, the proposed-likelihood function (their Equation 12) differs from that presented by Thompson (1975) and generalized here in Table 1. The following illustrates this incompatibility beginning with the compact notation used by Ritland (2000), which expresses as a single expression the likelihood function for samples from a noninbred population (Thompson 1975).

First, let

$$
\begin{aligned}
A &= \frac{1}{2}[\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}] \\
&= \frac{1}{2}[S_{ac}S_{bd} + S_{ad}S_{bc}] \qquad \text{(B1)}
\end{aligned}
$$

$$
\begin{aligned}
B &= \frac{1}{4}[(\delta_{ik} + \delta_{jk})p_l + (\delta_{il} + \delta_{jl})p_k] \\
&= \frac{1}{4}[(S_{ac} + S_{bc})p_d + (S_{ad} + S_{bd})p_c] \qquad \text{(B2)}
\end{aligned}
$$

$$
\begin{aligned}
C &= p_k p_l \\
&= p_c p_d \qquad \text{(B3)}
\end{aligned}
$$

$$
\begin{aligned}
D^{-1} &= p_i p_j (2 - \delta_{ij})(2 - \delta_{kl}) \\
&= p_a p_b (2 - S_{ab})(2 - S_{cd}). \qquad \text{(B4)}
\end{aligned}
$$

For each of these pairs of terms, the first uses the notation of Ritland (2000) and the second uses the notation of Lynch and Ritland (1999). Letting L_R refer to the likelihood function given by Ritland (2000, Equation 3.4),

$$
DL_R = \Delta_7 A + \Delta_8 B + \Delta_9 C. \qquad \text{(B5)}
$$

Noting that Δ_{xy} and ϕ_{xy} of Lynch and Ritland (1999) correspond to Δ_7 and Δ_8 of Ritland (2000) and that $\Delta_7 + \Delta_8 + \Delta_9 = 1$, some algebra demonstrates that this is equivalent to

$$
DL_R = DL_{LR} + (2\Delta_7 - \Delta_8)(B - C), \qquad \text{(B6)}
$$

where L_{LR} is the likelihood function proposed by Lynch and Ritland (1999, Equation 12). Because $B - C = 0$ is never true, the two likelihood functions are not equivalent.