Molecular Ecology (1998) **7**, 639–655

Statistical confidence for likelihood-based paternity inference in natural populations

T. C. MARSHALL, J. SLATE, L. E. B. KRUUK and J. M. PEMBERTON
Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh, EH9 3JT, UK

Abstract

Paternity inference using highly polymorphic codominant markers is becoming common in the study of natural populations. However, multiple males are often found to be genetically compatible with each offspring tested, even when the probability of excluding an unrelated male is high. While various methods exist for evaluating the likelihood of paternity of each nonexcluded male, interpreting these likelihoods has hitherto been difficult, and no method takes account of the incomplete sampling and error-prone genetic data typical of large-scale studies of natural systems. We derive likelihood ratios for paternity inference with codominant markers taking account of typing error, and define a statistic Δ for resolving paternity. Using allele frequencies from the study population in question, a simulation program generates criteria for Δ that permit assignment of paternity to the most likely male with a known level of statistical confidence. The simulation takes account of the number of candidate males, the proportion of males that are sampled and gaps and errors in genetic data. We explore the potentially confounding effect of relatives and show that the method is robust to their presence under commonly encountered conditions. The method is demonstrated using genetic data from the intensively studied red deer (*Cervus elaphus*) population on the island of Rum, Scotland. The Windows-based computer program, CERVUS†, described in this study is available from the authors. CERVUS can be used to calculate allele frequencies, run simulations and perform parentage analysis using data from all types of codominant markers.

Keywords: CERVUS, likelihood, LOD score, microsatellites, paternity inference, red deer

Received 19 October 1997; revision received 24 November 1997; accepted 1 December 1997

Introduction

In classical paternity inference one excludes as many as possible of the candidate males from paternity of a particular offspring using the available genetic data. If this procedure yields a single nonexcluded male, paternity is assigned to that male. This is the basic methodology underlying human parentage testing (Chakraborty *et al*. 1974), and there are several examples of the approach in wild animal populations (e.g. Morin *et al*. 1994; Hogg & Forbes 1997; Keane *et al*. 1997). However, exclusion by itself may be insufficient to unambiguously resolve paternity in a considerable proportion of paternity tests, even using a series of very polymorphic codominant markers where the probability of excluding an arbitrary unrelated male is very high (Chakraborty *et al*. 1988). While exclusion may be a useful starting point in paternity inference, a statistically based method is needed to assign paternity when multiple males are nonexcluded. The use of likelihood (Edwards 1972) for inference of relationships using genetic data was first explored in detail by Thompson (1975, 1976a), who showed that likelihood is an efficient approach to the evaluation of alternative relationships between a given pair of individuals in inference of human

Correspondence: T. C. Marshall. Fax: +44 (0)131650 6564; E-mail: tristan.marshall@ed.ac.uk.

†CERVUS 1.0 for Windows 95 is available from the authors. Starting with text-based genetic data, CERVUS determines allele frequencies, performs simulations and analysis paternity in study populations within a single easy-to-use Windows framework, with comprehensive on-line help. CERVUS can be downloaded from http://helios.bto.ed.ac.uk/evolgen/ or can be obtained by E-mailing the corresponding author. If neither of these methods is possible, a 3.5" floppy disk can be sent to the corresponding author. Note that not all the analyses presented in this paper can be carried out with the release version of the program.

pedigrees. However, the real problem of paternity inference in polygynous natural populations is subtly different. At issue is the evaluation of alternative pairs of individuals for a given relationship (father–offspring).

Meagher (1986) developed the likelihood approach of Thompson when he analysed allozyme data from a natural population of the lily *Chamaelirium luteum*, awarding paternity to the male with the highest log-likelihood ratio or LOD score (the likelihood ratio is the likelihood of paternity of a particular male relative to the likelihood of paternity of an arbitrary male). If two or more males were equally likely (usually because their genotypes were identical), paternity was left unassigned. However, Meagher did not assess the statistical confidence in the paternities that were awarded. Foltz & Hoogland (1981) carried out an analysis of paternity in prairie dogs (*Cynomys ludovicianus*), also using allozyme data. In each litter, to evaluate the likelihood of paternity for a series of males, Foltz and Hoogland used the difference in the log-likelihoods of the most-likely resident male and the most-likely nonresident male (ΔL), and chose an arbitrary threshold of four as a criterion for awarding paternity to a nonresident male. As in Meagher's (1986) study, Foltz and Hoogland were not able to assess the significance of their ΔL values, and they made a priori assumptions about the relative likelihood of paternity of resident and nonresident males. Furthermore, Meagher (1986) criticised ΔL on the basis that it is not a valid likelihood ratio, because one hypothesis is not nested within the other.

An alternative to categorical assignment of paternity to the most-likely male is to assign paternity fractionally to all nonexcluded males based on their relative likelihoods of paternity (Devlin *et al*. 1988; Roeder *et al*. 1989; Smouse & Meagher 1994). This approach allows population-level patterns of paternity to be assessed, even when the discriminatory power of marker loci is low. Smouse & Meagher (1994) compared a maximum likelihood fractional method against the earlier categorical analysis of Meagher (1986) for the same *Chamaelirium* data set. The correlation between individual male success as measured by the two methods was significant but weak ($r = 0.32$). Smouse and Meagher argue that the fractional method is superior because it uses all the available data. However, the method will systematically underestimate variance in male reproductive success (Devlin *et al*. 1988; Smouse & Meagher 1994). We have therefore concentrated on the most-likely or categorical approach to paternity assignment, which has much improved power with highly polymorphic DNA markers such as microsatellites, and has the additional advantage that the resulting paternity data can be used for analyses that require individual parent–offspring links, such as the calculation of inbreeding coefficients and heritabilities.

Published work on nonhuman paternity inference with codominant markers does not deal with the problems encountered when analysing the data typically obtained in large-scale genetic studies of natural populations. Paternity inference may be carried out with or without maternal genetic data, possibly within the same study; the two situations are statistically distinct and should be analysed appropriately. Statistical analysis should take account of the number of candidate males and be able to resolve paternity with confidence when not all candidate males are sampled. At the genotypic level, individuals may not be typed at every locus and, perhaps most importantly, the analysis should be robust to errors in typing. Nearly all published studies of natural populations treat a mismatch between a male and a putative offspring as conclusive evidence for exclusion of that male from paternity. In practice, a mismatch could result either from a genuine nonrelationship or from a laboratory typing error, a reality acknowledged for some time in human pedigree analysis (Thompson 1976b; Ashton 1980; Lathrop *et al*. 1983). When microsatellite markers are used, mutations (Queller *et al*. 1993) and null alleles (Callen *et al*. 1993; Phillips *et al*. 1993; Pemberton *et al*. 1995) may also generate mismatches between genuine relatives at measurable frequencies.

In this paper, we develop the likelihood-based approach of Thompson (1975, 1976a) and Meagher (1986). Paternity is assigned to a particular male if the likelihood ratio is large relative to the likelihood ratios of alternative males. Similar to Meagher (1986), we express likelihood ratios as LOD scores (i.e. the logarithm of the likelihood ratio). Unfortunately, significance levels for LOD scores cannot conveniently be derived analytically (Edwards 1972; Meagher 1986). However, Thompson & Meagher (1987) showed that the ratio of the likelihood ratios of two males is a true likelihood comparison of alternative father–offspring relationships. Considering the two most-likely males, we define the logarithm of the ratio of likelihood ratios to be Δ (equal to the difference in LOD scores), and use computer simulation of paternity inference to generate criteria for Δ appropriate for paternity assignment in the study population. The use of simulation of paternity inference in order to evaluate the significance of LOD scores was also suggested by Taylor *et al*. (1997).

A statistical pitfall of using conventional likelihood ratios to assess paternity has been pointed out by Thompson (1976a, 1976b) and Thompson & Meagher (1987). If there exist in the population full sibs of the offspring whose paternity is being tested, and no genetic data are available from the mother, nonexcluded full sibs on average have a higher likelihood of paternity than the true father. We examine the effect of relatives of the offspring on paternity inference with and without maternal genetic data, and compare this with the effect of relatives of the true father.

We also demonstrate our approach using genetic data from the intensively studied red deer (*Cervus elaphus*) population on the island of Rum, Scotland. For convenience, we refer throughout to paternity inference; however, our system is equally applicable to inference of maternity (for example Jones & Avise 1997).

Materials and methods

Likelihood in paternity testing

Likelihood analysis takes data as a starting point, and evaluates hypotheses given that data (Edwards 1972). The likelihood L of a hypothesis H given data D can be written $L(H|D)$. The likelihood of one hypothesis (e.g. H_1) is always evaluated relative to another (e.g. H_2). This is the likelihood ratio, written as $L(H_1,H_2|D)$:

$$L(H_1, H_2|D) = \frac{P(D|H_1)}{P(D|H_2)} \quad (1)$$

where $P(D|H_i)$ is the probability of obtaining data D under hypothesis H_i. In the context of paternity inference, the data D are the genotypes of offspring, mother, and alleged father at a particular locus. The hypothesis of interest H_1 is that the alleged father is the true father, and this is tested against hypothesis H_2 that the alleged father is an unrelated individual selected at random from the population. The following interpretation (eqn 2 – eqn 4) is based on that of Meagher (1986), and assumes that the mother's genotype is known.

Let g_m, g_a and g_o represent the genotypes of mother, alleged father and offspring, respectively, at a given locus. The likelihood that the mother and alleged father are the parents of the offspring can then be expressed:

$$L(H_1|g_m,g_a,g_o) = T(g_o|g_m,g_a).P(g_m).P(g_a) \quad (2)$$

Here, $T(g_o|g_m,g_a)$, the probability of the offspring's genotype given the genotypes of the mother and alleged father, is the Mendelian segregation or transition probability. $P(g_m)$ and $P(g_a)$ are the frequencies of the mother's and alleged father's genotypes in the population. The likelihood that the mother is the parent of the offspring and the father is a randomly chosen individual from the population is expressed:

$$L(H_2|g_m,g_a,g_o) = T(g_o|g_m).P(g_m).P(g_a) \quad (3)$$

where $T(g_o|g_m)$ is the probability of the offspring's genotype given the mother's genotype. The likelihood ratio (eqn 2 divided by eqn 3), represents how much more likely it is that the alleged father, rather than an arbitrary male, passed his genes to the offspring (Aitkin 1995).

$$L(H_1, H_2|g_m,g_a,g_o) = \frac{T(g_o|g_m,g_a).P(g_m).P(g_a)}{T(g_o|g_m).P(g_m).P(g_a)} = \frac{T(g_o|g_m,g_a)}{T(g_o|g_m)} \quad (4)$$

The likelihood ratios for all compatible genotypic combinations at a codominant autosomal locus are presented in Table 1. (Note that eqn 4 is referred to as the Paternity Index (PI) in human paternity testing (Pena & Chakraborty 1994)).

In cases where the mother's genotype is unknown, the likelihood ratio is different:

$$L(H_1, H_2|g_a,g_o) = \frac{T(g_o|g_a).P(g_a)}{P(g_o).P(g_a)} = \frac{T(g_o|g_a)}{P(g_o)} \quad (5)$$

Here, the denominator, $P(g_o)$, is the frequency of the offspring's genotype. Allele frequencies may only be used to estimate genotype frequencies if Hardy–Weinberg

Table 1 Likelihood ratios for all compatible mother–alleged father–offspring trios. *X* represents any allele other than *B*; *Y* represents any allele that is neither *B* nor *C*. The frequencies of alleles *B* and *C* are denoted *b* and *c*. The likelihood ratio, $L(H_1,H_2)$, is the probability of the offspring's genotype given the mother's and alleged father's genotypes, $T(g_o|g_m,g_a)$, divided by the probability of the offspring's genotype given the mother's genotype, $T(g_o|g_m)$. A similar table is shown in more condensed form in Brenner (1997)

Offspring's genotype (g_o)	Alleged father's genotype (g_a)	Mother's genotype (g_m)	$T(g_o\|g_m,g_a)$	$T(g_o\|g_m)$	$L(H_1,H_2)$
BB	*BB*	*BB*	1	b	$1/b$
BB	*BX*	*BB*	1/2	b	$1/2b$
BB	*BB*	*BX*	1/2	$b/2$	$1/b$
BB	*BX*	*BX*	1/4	$b/2$	$1/2b$
BC	*BB*	*CC*	1	b	$1/b$
BC	*BX*	*CC*	1/2	b	$1/2b$
BC	*BB*	*CY*	1/2	$b/2$	$1/b$
BC	*BX*	*CY*	1/4	$b/2$	$1/2b$
BC	*BB*	*BC*	1/2	$(b+c)/2$	$1/(b+c)$
BC	*BY*	*BC*	1/4	$(b+c)/2$	$1/2(b+c)$
BC	*BC*	*BC*	1/2	$(b+c)/2$	$1/(b+c)$

equilibrium holds. The likelihood ratios for paternity inference without maternal genetic information, assuming Hardy–Weinberg equilibrium, are presented in Table 2.

When several unlinked marker loci are used in paternity inference, the likelihood ratios derived at each locus may be multiplied together and the natural ($\log_e$) logarithm taken. Meagher (1986) terms the natural logarithm of the combined likelihood ratio the LOD score. [Meagher's definition of the LOD score differs from that used in genetic mapping, where the LOD score is defined as the common ($\log_{10}$) logarithm of the likelihood ratio. For consistency with standard likelihood analysis (Edwards 1972), and with previous work in paternity inference (e.g. Thompson 1975), we follow Meagher's (1986) definition.] LOD score of zero implies that the alleged father is equally as likely to be the father of the offspring as a randomly selected male. A positive LOD score implies that the alleged father is more likely to be the father of the offspring than a randomly selected male; negative LOD scores may occur if the alleged father and offspring share a particularly common set of alleles.

Mismatches and typing errors

If genetic data are perfect, a mismatch at a single locus between alleged father and offspring can be logically treated as a paternity exclusion. However, data are often not perfect, so it is unwise to exclude males entirely from paternity on this basis. A mismatch can result not only from nonpaternity but also from erroneously typed paternal, maternal or offspring genotypes, or from marker mutation or null alleles (see the Discussion). The use of a likelihood approach allows us to introduce a parameter, the error rate, which takes account of potential imperfections in the data. We define an error to be the replacement of the true genotype at a particular locus in an individual with a random genotype. Errors may occur in the genotypes of offspring, mother or alleged father, or in some combination of the three. The derivations of likelihood ratios incorporating errors are shown in Appendix 1. These likelihood ratios are used, assuming Hardy–Weinburg equilibrium, for all analyses presented in this paper.

Assignment of paternity using LOD scores

In order to discriminate between nonexcluded males, we define a statistic Δ as the difference in LOD scores between the most-likely male and the next most-likely male. Let n be the number of candidate males with a LOD score greater than zero. The LOD score of male i is denoted LOD_i, and the males are ranked such that $LOD_i \geq LOD_{i+1}$ for $1 \leq i < n$ (so, for example, the LOD score of the most-likely male is denoted LOD_1). Then Δ is defined as follows:

$n \geq 2$, $\Delta = LOD_1 - LOD_2$
$n = 1$, $\Delta = LOD_1$
$n = 0$, Δ undefined

Without a threshold LOD score of zero, Δ is sensitive to LOD_2. If LOD_2 is very negative (typically when all candidate males except the most-likely male mismatch the offspring at several loci), Δ is large whatever the value of LOD_1. A threshold LOD score of zero stabilizes Δ because it always lies between zero and LOD_1.

Simulation of paternity inference

We use simulations to assess the significance of Δ values. The simulation analysis of the program CERVUS emulates the steps of paternity inference using allele frequencies at loci screened in a given study population (Fig. 1). Parallel simulations are carried out for paternity inference with and without maternal genetic data.

Offspring's genotype (g_o)	Alleged father's genotype (g_a)	$T(g_o \mid g_a)$	$P(g_o)$	$L(H_1,H_2)$
BB	*BB*	b	b^2	$1/b$
BB	*BX*	$b/2$	b^2	$1/2b$
BC	*BB*	c	$2bc$	$1/2b$
BC	*BY*	$c/2$	$2bc$	$1/4b$
BC	*BC*	$(b+c)/2$	$2bc$	$(b+c)/4bc$

Table 2 Likelihood ratios for all compatible alleged father–offspring pairs, in the absence of a genotyped mother. *X* represents any allele other than *B*; *Y* represents any allele that is neither *B* nor *C*. The frequencies of alleles *B* and *C* are denoted b and c. Likelihood ratios are calculated on the basis that Hardy–Weinberg equilibrium holds. The likelihood ratio, $L(H_1,H_2)$, is the probability of the offspring's genotype given the alleged father's genotype, $T(g_o \mid g_a)$, divided by the probability of the offspring's genotype, $P(g_o)$. A similar table is shown in more condensed form in Brenner (1997)

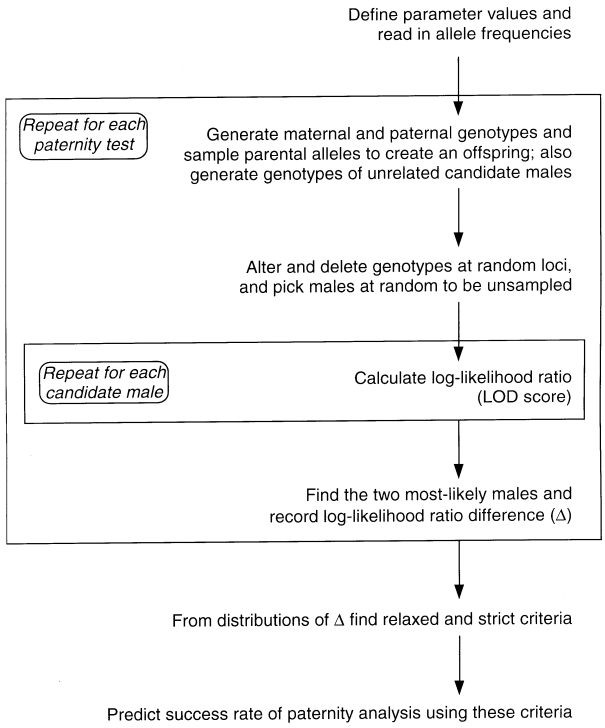

Fig. 1 A flow chart illustrating the operation of the CERVUS program's simulation of paternity inference. Boxes indicate repeated loops. Simulation of paternity inference where mothers are unsampled is carried out in a parallel simulation.

Assuming Hardy–Weinberg equilibrium, a maternal genotype and a paternal genotype are generated from allele frequencies observed in the study population, and an offspring genotype is derived by Mendelian sampling of the parental alleles. Genotypes are also generated for a number of unrelated candidate males. The genotypic data for all individuals are then altered to reflect the existence of unsampled males, missing loci and incorrectly typed loci, according to the values of the parameters described below. Next, each candidate male is considered in turn as the alleged father, beginning with the true father. LOD scores are calculated for all males for whom genetic data exist. Once all males have been considered, the most-likely and second-most-likely males are identified and the Δ score calculated (all males with LOD scores of 0 or less are ignored). The value of Δ is recorded along with the status of the most-likely male (i.e. whether or not this is the true father).

Genetic data are generated and paternity tests carried out for a large number of simulated offspring in order to generate distributions of Δ (Fig. 2). A total of 10 000 tests is sufficient in most cases. The following parameters (Table 3) are included in the simulation model in order to make simulated genetic data realistic.

Number of candidate males. The average number of males that are candidates for paternity of each offspring. The number of candidate males can be estimated from field data, and includes males that are not sampled. Candidate males other than the true father are assumed to be unrelated to the mother-father-offspring trio (the effect of relaxing this assumption is explored in the Results).

Proportion of candidate males sampled. The average fraction of candidate males for whom genotypic data are available. The proportion of sampled males can be estimated from field data. The true father falls into this category with the same probability as any other candidate male.

Proportion of loci typed. The fraction of loci typed, averaged across all loci and individuals. Missing genotypes are scattered at random across loci and individuals, including the mother, true father and offspring. The proportion of loci typed can readily be calculated from the genetic data used to estimate allele frequencies.

Error rate. The fraction of loci typed incorrectly, averaged across all loci and individuals. An error is defined as the replacement of the true genotype at a given locus with a genotype selected at random under Hardy–Weinberg assumptions. Under this definition, an erroneous genotype will sometimes be the same as the true genotype. If mother–offspring pairs are known from field data, the error rate can be estimated from the frequency of mismatches (i.e. no alleles in common) between mothers and their offspring, bearing in mind that as defined not all errors alter the observed genotype, and that not all alterations of genotype can be detected as mother–offspring mismatches (see Appendix 2). Note that the error rate parameter is used in the simulation of paternity inference both to generate simulated genetic data and in the calculation of likelihoods (eqn 6 – eqn 9).

Identifying critical values of Δ

The final stage of the simulation is to find critical values of Δ so that the significance of Δ values found in paternity inference in the study population can be tested. The program compares the distribution of Δ scores for cases where the most-likely male was the true father with that for cases where the most-likely male was not the true father (Fig. 2). Assuming, as an example, that a criterion is required for Δ which gives 95% confidence, the program identifies the value of Δ such that 95% of Δ scores exceeding this value are obtained by true fathers. If the program fails to find such a value of Δ (typically because the resolving power of the markers is insufficient), the critical value of Δ is set to an arbitrary high value of 99.99. When a male fulfilling the 95% confidence criterion is

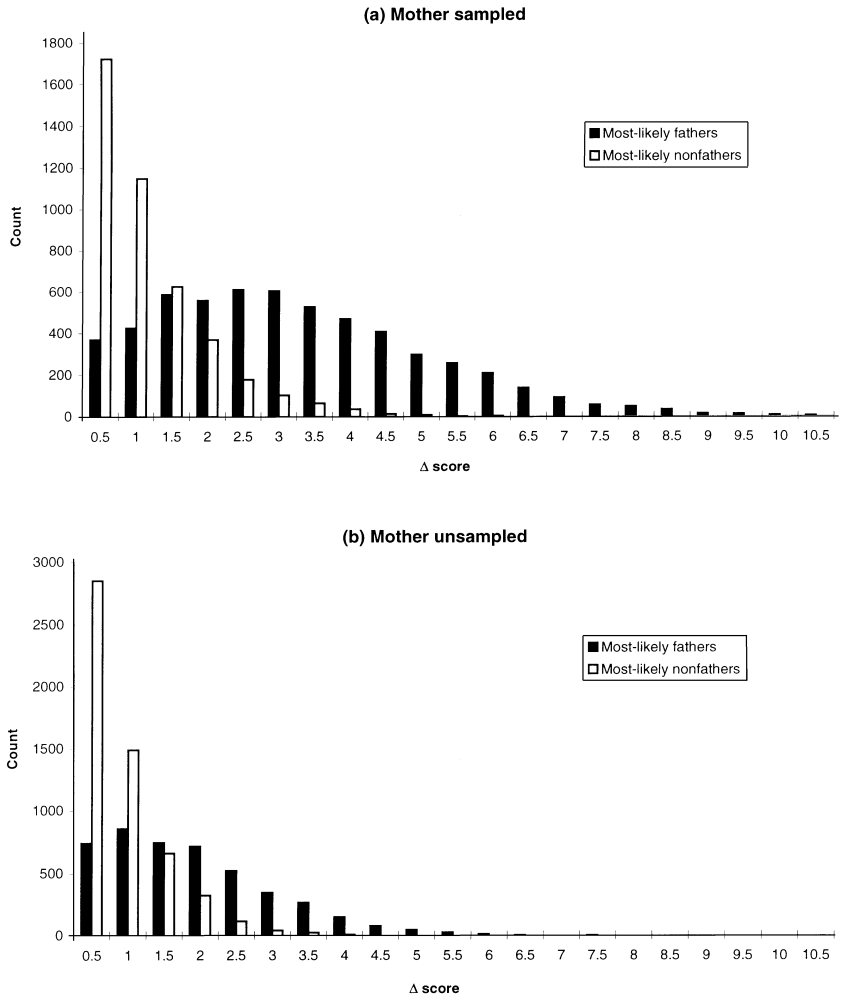

Fig. 2 Histograms of Δ scores generated by simulation of paternity inference using allele frequencies from Rum red deer and the parameter values shown in Table 3. In each plot the histogram of N_t cases where the most-likely male is the true father (filled bars) is interleaved with the histogram of N_u cases where the most-likely male is not the true father (white bars). Critical values of Δ are calculated from the degree of overlap of the two distributions shown in each plot. Critical values for 80% and 95% confidence are listed in Table 5. Plot (a) is from a simulation when the mother was sampled, and plot (b) is from a simulation when the mother was unsampled. Δ categories are labelled with their upper limits.

assigned paternity of an offspring, we describe the father–offspring relationship as a 95% confident paternity.

Confidence levels can be thought of as levels of tolerance of 'false-positive' paternities, or paternities assigned to males who match by chance (Type I Error). For some purposes the number of paternities obtained may be most important, while in other situations very accurate paternity assignment may be required. The program therefore calculates separately relaxed and strict confidence levels. In this paper we interpret relaxed confidence as 80% and strict confidence as 95%. For each level of confidence, the program shows the percentage of simulated paternity tests in which the Δ score of the most-likely male exceeded the critical value of Δ (i.e. the percentage of tests in which paternity was assigned), a statistic we refer to as the success rate.

Red deer on Rum

The unmanaged red deer population in the North Block of the island of Rum (Inner Hebrides, Scotland) has been the focus of intensive study since 1971. Detailed descriptions of the study site are available in Clutton-Brock *et al.* (1982) and Clutton-Brock & Albon (1989). A total of 1168 individually monitored deer have been typed at up to three protein and nine microsatellite loci (details in

Table 3 The parameters used in simulation of paternity inference with the CERVUS program, and the values used in simulations for Rum red deer presented in this paper. The number of candidate males and the proportion of males sampled are average values from the ruts between 1981 and 1995 that gave rise to the calves born 1982–96. The proportion of loci typed and the error rate are average values across the 12 loci screened. See text for details on choice of parameter values

Parameter	Value used
Number of candidate males	75
Proportion of candidate males sampled	0.65
Proportion of loci typed	0.854
Rate of typing error	0.01
Number of tests	10,000
Relaxed confidence level	80%
Strict confidence level	95%

Table 4 The three protein loci (top) and nine microsatellite loci (bottom) used to generate the simulation data presented in this paper. Oar and MAF microsatellites are randomly cloned sequences from domestic sheep (*Ovis aries*); Cel microsatellites are randomly cloned sequences from Rum red deer (*Cervus elaphus*); TGLA94 is a randomly cloned microsatellite sequence from domestic cattle (*Bos taurus*). Expected heterozygosity was calculated using eqn 8.4 of Nei (1987)

Locus name	Number of alleles	Number of genotypes	Expected heterozygosity	Reference
Mannose phosphate isomerase	2	890	0.257	Pemberton *et al*. (1988)
Isocitrate dehydrogenase	2	893	0.500	Pemberton *et al*. (1988)
Transferrin	2	702	0.428	Pemberton *et al*. (1988)
OarFCB193	11	1044	0.761	Buchanan & Crawford (1993)
OarFCB304	9	1102	0.789	Buchanan & Crawford (1993)
CelJP15	10	1107	0.837	J. Pemberton, unpublished data
CelJP27	6	1059	0.691	J. Pemberton, unpublished data
CelJP38	8	999	0.786	J. Pemberton, unpublished data
MAF35	7	994	0.673	Swarbrick *et al*. (1991)
MAF109	6	1056	0.751	Swarbrick & Crawford (1992)
OarCP26	13	1086	0.722	Ede *et al*. (1995)
TGLA94	9	1042	0.808	Georges & Massey (1992)

Table 4), and individuals were on average typed at 85% of loci. Protein genotyping protocols are described in Pemberton *et al*. (1988), and microsatellite genotyping procedures are the same as those described by Bancroft *et al*. (1995) for Soay sheep. Genotypic frequencies at all loci conform to Hardy–Weinberg expectations.

The multiplication of likelihood ratios over several loci assumes that the loci segregate independently. This assumption is violated if any pair of loci is in linkage disequilibrium. For example, high values of linkage disequilibrium are likely for a series of loci known to lie within a single gene cluster (e.g. the Major Histocompatibility Complex) or in zones of hybridization between two populations, subspecies or species (Barton & Gale 1993). In red deer, we are aware that two pairs of the microsatellite markers used in our analysis each share a linkage group. However, neither pair is thought to be tightly linked (M. Tate, personal communication), and loose linkage of one or two pairs of markers is not thought to seriously bias multilocus likelihood calculations (Meagher 1986). When calculating LOD scores, we therefore treat all loci as if they segregate independently (i.e. we assume that all pairs of markers are in linkage equilibrium).

Allele frequencies from the total Rum sample were used for all the simulation results presented below. Paternity was analysed for a subset of 875 calves born between 1982 and 1996. Red deer have an annual rutting season in the autumn, and single calves are born the following spring. In each year, the candidate males for paternity of each calf are all males classified as behaviourally active at any time during the rut preceding the birth of that calf. Typically this includes any male 3-years old or more observed in the study area at this time. Stags aged 2 years or less have never been observed to obtain matings, and in practice mating opportunities for stags less than 5-years old are extremely infrequent (F. Guinness, unpublished data). The paternity inference procedure is thus conservative, in that it considers any male potentially able to mate as a candidate male. In ruts between 1981 and 1995 (giving rise to the calf cohorts 1982–96), there was an average of 75 candidate males, of which 65% were sampled. These values were used in the simulation.

Where maternal genetic data are used in inferring paternity, maternity is based on field observations. Inaccuracies in assignment of maternity are likely to show up as mismatches across multiple loci, and no such cases have been found (Pemberton *et al*. 1988, 1992; our unpublished data). The sporadic mismatches between maternal and offspring genotypes that do occur are interpreted as errors in the typing process. Our estimate for the overall rate of typing error, given that many errors go undetected, is 1% (see Appendix 2).

Results

In this section, we explore the results generated by the CERVUS program. We examine in turn the statistical properties of Δ, the repeatability of simulation results, the importance of typing errors, the effect of varying simulation parameters on the predicted resolving power of markers and the impact of relatives on paternity inference. Finally,

we apply critical Δ scores from the simulation to the red deer data set. All simulations use the allele frequencies for Rum red deer screened at the loci described in Table 4.

Distribution of Δ

The distributions of Δ scores, from a simulation of paternity inference using the parameter values listed in Table 3, are shown in Fig. 2. In general, pairs of distributions of Δ for true fathers and nonfathers tend to fall into three categories, depending on the ratio of the number of true fathers, N_t (filled bars in Fig. 2), to nonfathers, N_u (white bars):

1 $N_t > N_u$. The critical value is low, and success rate is high (e.g. Fig. 2a). In the special case of $N_t/(N_t + N_u)$ being greater than or equal to the confidence level, the critical value is set to zero.
2 $N_t \approx N_u$. The critical value takes an intermediate value, but success rate is heavily dependent on the degree of overlap of the two distributions. If the overlap is narrow, more paternities can be assigned than if the overlap is wide.
3 $N_t < N_u$. The critical value is high, and the success rate is low (e.g. Fig. 2b).

Which of these three scenarios occurs varies according to whether or not the mother is sampled, the allele frequencies and the parameter values used in the simulation. $N_t + N_u$ may be less than the total number of paternity tests carried out, as under certain parameter conditions there may be a proportion of tests where no male has a LOD score greater than zero.

The distributions shown in Fig. 2 were used to derive critical Δ scores and the corresponding predicted success rates (Table 5). Δ criteria were larger when mothers were not sampled and larger for higher confidence. Success rates were smaller when mothers were not sampled, and smaller for higher confidence.

Repeatability of simulation results

The simulation generates repeatable results with 10 000 paternity tests. More than 95% of critical Δ values were within 10% of the mean critical Δ value, based on 16 runs using the parameter values shown in Table 3. From the same runs, 95% of predicted success rates were within 2% (expressed as a percentage of paternity tests resolved) of the mean predicted success rate. In other words, the variation in critical Δ values that did occur did not lead to large variations in predicted success rates. Critical Δ values at the lower end of the range did not give rise to many Type I Errors (incorrect assignment of paternity), and critical Δ values at the upper end of the range did not give rise to many Type II Errors (failure to assign paternity).

The ranges of critical Δ values and predicted success rates described applied to 80% and 95% confidence, and to simulations with and without mothers sampled. In practice, the ranges were a little narrower at 80% confidence. If a very high level of accuracy is required, or if very high levels of confidence are demanded (e.g. 99% or more), simulations with more than 10 000 paternity tests may be carried out at a cost of increased computing time.

In the following sections describing simulation results, we concentrate on the comparison between paternity inference with and without sampled mothers at 80% confidence. In all cases, similar patterns were observed at 95% confidence.

Importance of typing errors

An error rate of zero ensures that any mismatch is treated as a paternity exclusion, and the likelihood equations then reduce to eqn 4 and eqn 5. Given that laboratory data are rarely error free, is it satisfactory or prudent to ignore errors in paternity inference? We examined the impact of errors on confidence, and whether taking account of errors alters success rates.

If simulations and paternity inference in the study population are carried out on the basis that genetic data are error free, Fig. 3 shows the true confidence in paternities allegedly assigned at 80% confidence for simulated genetic data including typing errors at various rates. In all cases, assuming genetic data to be error free led to overestimates of the confidence in paternities assigned when data include errors. For example, when mothers were unsampled and the true error rate was 2%, paternities assigned

Simulation results (n = 10 000)	Mother sampled		Mother unsampled	
	80%	95%	80%	95%
Critical value of Δ	1.13	2.78	1.49	3.15
Proportion of paternities	59.30%	28.86%	25.33%	4.35%

Table 5 The critical Δ scores and number of red deer paternity tests predicted to be resolved by simulation using the program CERVUS. Parallel simulations were carried out for paternity inference with sampled mothers and unsampled mothers. Relaxed (80% confidence) and strict (95% confidence) criteria are shown, along with the proportion of paternity tests (of 10 000) in which a male fulfilled the required criterion (i.e. was awarded paternity)

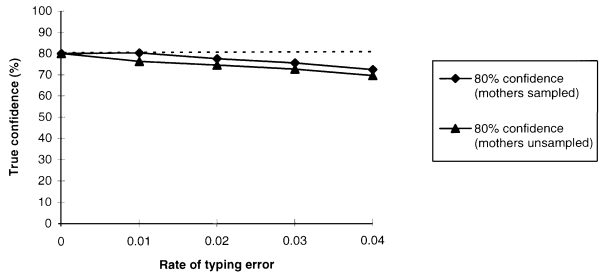

Fig. 3 The effect of ignoring typing error on confidence of paternity inference with error-prone genetic data. Simulation criteria from a simulation with an error rate of zero were applied to simulated data containing errors at rates from 1 to 4%; likelihood calculations in all cases used an error rate of zero. The true confidence was calculated as the ratio $N_t/(N_t + N_u)$. The dotted line represents the null hypothesis of no effect of ignoring typing error on confidence. Values used for other simulation parameters are shown in Table 3.

with an apparent confidence of 80% (based on simulations assuming no errors) had a true confidence of 74%.

Success rate may be improved by allowing for errors, because a true father that was previously excluded on account of typing error may now have a LOD score sufficient to be identified as the father. On the other hand, allowing for typing errors may mean that unrelated candidate males that were previously excluded on account of mismatches at just one or two loci may now have similar LOD scores to the true father, meaning that the LOD score of the true father is no longer sufficiently large to award paternity. The first of these effects predominates when there is redundancy of information across marker loci, while the second effect predominates when the resolving power of markers is limited. Figure 3 suggests that at least some of the additional paternities that may be awarded by ignoring errors are of inferior confidence. For example, '80% confident' paternities that were awarded by ignoring errors were not necessarily secure at 80% confidence when the correct simulation criteria were applied.

Success of paternity inference

The responses of success rates to parameter changes are shown in Fig. 4, expressed as a percentage of paternity tests resolved with 80% confidence. In each case one parameter was varied while the others were held constant at the values shown in Table 3. Simulations were carried out assuming 65% of males were sampled (except Fig. 4b where this parameter itself was varied). Thus success rates were unlikely to greatly exceed 65%, though false-positive paternities, especially at lower levels of confidence, can push the percentage a little higher.

Number of candidate males. Figure 4a shows the effect of varying the number of candidate males. Fewer paternity

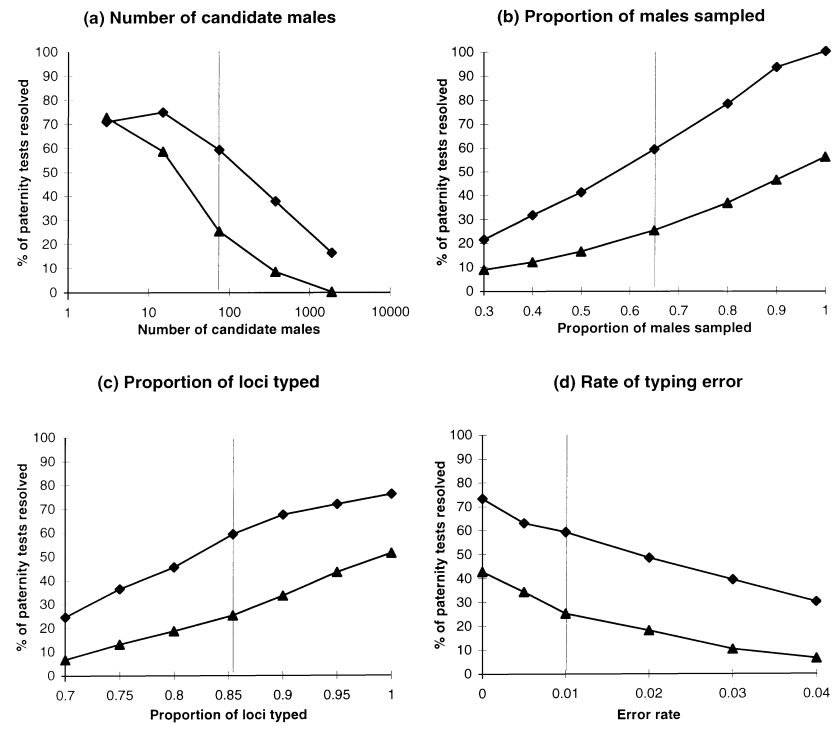

Fig. 4 The effect on the predicted success of paternity inference on varying the main parameters included in the model. (a) number of candidate males, (b) proportion of males sampled, (c) proportion of loci typed, and (d) rate of typing error. In each case, the percentage of paternity tests resolved at 80% confidence is shown, with all other parameters being held at the values in Table 3. The values used for simulation of the real Rum red deer data are indicated by a vertical line. Graph (a) uses a log scale on the *x*-axis. The legend is as for Fig. 3.

tests were resolved as the number of candidate males increased. Similar success rates were obtained when choosing between 100 candidate males with sampled mothers and choosing between 10 candidate males when mothers were unsampled, and this ratio is maintained across most of the chart. If the number of candidate males was large, maternal genetic data were essential to assignment of paternity with confidence of 80% using the set of markers screened in Rum red deer.

Proportion of candidate males sampled. Figure 4b shows the effect of varying the proportion of candidate males sampled. Success rates increased as the proportion of sampled candidates males increases.

Proportion of loci typed. Figure 4c shows the effect of varying the proportion of loci typed. The more complete the genetic data, the greater the success rates.

Rate of typing error. Figure 4d shows the effect of varying the rate of typing error. Fewer paternity tests were resolved as the error rate increased. Note that the error rate is included both in the generation of simulated genetic data and in the calculation of likelihoods.

The effect of relatives

Thus far, simulations have assumed that candidate males other than the true father are unrelated to the mother–father–offspring trio. We explored the effect of introducing relatives into the pool of candidate males, examining their effect on confidence in paternity assignments made assuming no relatives were present. We compared the effect of the degree of relatedness when candidate males were related to the offspring with the effect when candidate males were related to the true parents. We also explored how the number of half sibs of the parents affects paternity inference.

Figure 5a shows the effect on confidence of introducing males related to the offspring, assessed by applying the Δ criteria in Table 5 to simulated paternity tests including five related males. Confidence in paternities assigned using these criteria declined with increasing relatedness. When mothers were sampled, the overestimate in confidence was small unless full sibs of the offspring ($r = 0.5$) were among the candidate males. When mothers were unsampled, the overestimate in confidence was large if males related as half sibs or more to the offspring ($r \geq 0.25$) were among the candidate males, suggesting that paternity cannot be resolved between the true father and the father's sons, especially sons by the same female. This situation was studied by Thompson & Meagher (1987). Whether or not mothers were sampled, success rates declined only slowly with increasing relatedness (data not shown). In other words, as relatedness increases, a similar number of paternities can be assigned but with reduced confidence.

A more common relatedness problem is posed by the presence of male relatives of the true father. Figure 5b shows the effect on confidence when five males related to the true father were among the candidate males, assessed as for Fig. 5a. Whether or not mothers were sampled, confidence declined only slowly with increasing relatedness, and even with five full sibs among the candidate males, true confidence was 72% when 80% confidence criteria from Table 5 was applied. Again, success rates declined only slowly with increasing relatedness (data not shown). Although we only considered relatives of the true father, note that when mothers are unsampled, male relatives of the mother have the same confounding effect on paternity inference as male relatives of the true father.

There is a straightforward explanation for why the impact of relatives of the offspring differs from the impact of relatives of the true father. Assuming no inbreeding, a full sib of the offspring and a full sib of the true father are both related to the true father with $r = 0.5$; a full sib of the

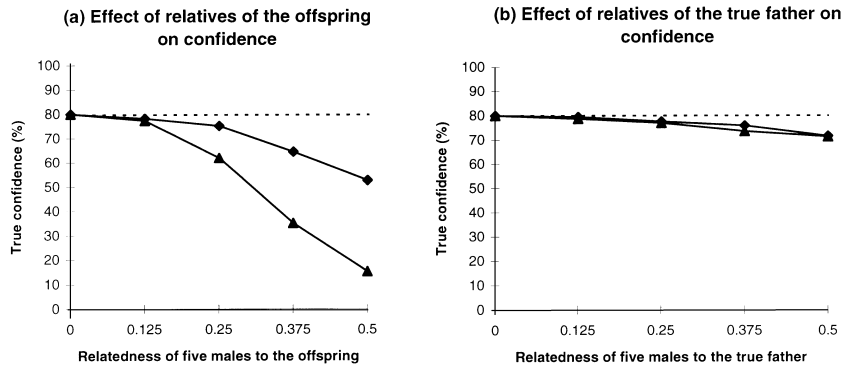

Fig. 5 The effect of including relatives among the candidate males on true confidence of parentage assignments, varying (a) relatedness of five males to the offspring, and (b) relatedness of five males to the true father. Simulation criteria from a simulation without relatives (Table 5) were applied to simulated genetic data containing five relatives of the specified level of relatedness. The true confidence was calculated as the ratio $N_t/(N_t + N_u)$. The dotted line represents the null hypothesis of no effect of relatives on confidence. Values used for simulation parameters are shown in Table 3. The legend is as for Fig. 3.

offspring is also related to the mother with $r = 0.5$ whereas a full sib of the true father is unrelated to the mother. If the mother is unsampled, a candidate male that is a full sib of the offspring has a probability of 0.75 of carrying either the offspring's paternal or maternal allele or both due to relatedness, whereas a candidate male that is a full sib of the true father has a probability of 0.5 of carrying the offspring's paternal allele due to relatedness, and cannot carry the offspring's maternal allele due to relatedness since the candidate male is unrelated to the mother. This explains why Fig. 5, a and b, differed when mothers were unsampled. If the mother is sampled, only the offspring's paternal alleles are used in the paternity test. Providing that the paternal allele can always be unambiguously identified, full sibs of the offspring and full sibs of the true father have the same impact on paternity inference at a given locus, both having a probability of 0.5 of carrying the offspring's paternal allele due to relatedness. In practice there are always some loci where the paternal allele cannot be unambiguously identified, either because of missing maternal genetic data, or because mother and offspring share the same heterozygous genotype (most likely for loci with few alleles). In these cases, a candidate male that is a full sib of the offspring has a higher probability of bearing at least one parental allele than a full sib of the true father for the same reason as when the mother is unsampled. This explains why confidence also declined more steeply in Fig. 5a than in Fig. 5b when mothers were sampled. Although some inbreeding is likely to occur in most natural populations, the qualitative conclusions of this argument are unchanged unless the level of inbreeding is very extreme.

In polygynous species, there may be many half sibs in each population, and half sibs of the true father may often be the father's closest relatives considered as candidate males in paternity inference. The likelihood system is insensitive to large numbers of half sibs of the true father. Even with 25 half sibs present (i.e. one third of all males), true confidence was 72% when 80% confidence criteria from Table 5 was applied (data not shown).

Paternity inference in Rum red deer

The critical values of Δ shown in Table 5 were applied in an analysis of paternity for 875 red deer calves born between 1982 and 1996 (Table 6), carried out using the CERVUS program. Paternity was determined with 80% confidence for 385 of the 655 calves tested with sampled mothers (59%). Less than half of the 80% confident paternities (28% of all tests) were secure at 95% confidence. For the 220 paternity tests where the mother was unsampled, only 67 (30%) gave an 80% confident paternity, and less than one third of these (9% of all tests) were secure at 95% confidence. More paternities could be assigned, or confidence in existing paternities increased, by sampling a larger number of candidate males, adding missing genetic data or retyping existing data to reduce the frequency of typing error (Fig. 4). An alternative strategy would be to type additional loci (our unpublished simulation data).

The success of paternity inference in Rum red deer is in close agreement with the predictions of the simulation. Success was high when mothers were sampled, but low when mothers were unsampled and, in both cases, fewer paternities were secure at 95% confidence than at 80% confidence. Differences could arise by cohort-by-cohort variation in number or sampling of candidate males, locus-by-locus variation in frequency of missing data or error rates, unequal distribution of reproductive success between sampled and unsampled males, inaccurate estimation of the error rate, and relatives of the parents amongst the candidate males. Despite these potentially confounding effects, the simulation appears to be a useful predictive tool.

Discussion

In this paper we extend standard likelihood-based paternity inference to deal with marker mistyping, and develop a likelihood-based statistic, Δ, for determining paternity using codominant molecular markers. Δ is tested against critical values derived by simulation of paternity inference using observed marker allele frequencies. The simulation

Number of paternities	Mother sampled $n = 655$		Mother unsampled $n = 220$	
	80%	95%	80%	95%
Observed	385 (59%)	184 (28%)	67 (30%)	19 (9%)
Expected	388 (59%)	189 (29%)	56 (25%)	10 (4%)

Table 6 The results of paternity inference using the program CERVUS for 875 Rum red deer calves born between 1982 and 1996. The criteria used to assign paternity are shown in Table 5. Cases where the mother was sampled were analysed separately from those where the mother was unsampled. The number of paternities obtained with 80% and 95% confidence (observed) are listed above the number of paternities predicted from success rates shown in Table 5 (expected)

system takes account of the number of candidate males, the proportion that is sampled, the completeness of genetic data and the rate of typing error in deriving these critical values.

The success of paternity inference using our approach is influenced by two major factors, apart from the number of candidate males and the quality of markers used. These factors are whether a sample is available from the mother, and the level of confidence required of paternities that are assigned. The results presented here show that there is a consistently large premium attached to obtaining the mother's genotype. Without this, at least 50% more loci are required to deliver a similar success rate at the same level of confidence (our unpublished simulation data). The results also show that there is a clear trade-off between the number of paternities assigned, and the confidence in those assignments.

It is important to realize that paternities assigned with 80% confidence are more accurate than can be achieved in most species by direct observation, and are also better than would be obtained in many studies by a purely exclusionary approach, where confidence in paternity of non-excluded males is generally unknown. Paternity inference studies that use a simple exclusionary approach can generate impressive probabilities of excluding an unrelated candidate male from paternity (e.g. Morin *et al*. 1994). However, these probabilities can give a misleading impression of the probability of paternity of a nonexcluded male. For the 385 paternities assigned with 80% confidence in Rum red deer for offspring with sampled mothers, the median value of the exclusion probability was 0.9998. Our analysis suggests that the number or polymorphism of microsatellite loci required to confidently (e.g. 95%) identify paternity may be rather higher than many laboratories, including our own, initially aim for.

Previously published methods for paternity inference in natural populations assume, often implicitly, that the entire pool of candidate males has been sampled (e.g. Meagher 1986; Devlin *et al*. 1988; Smouse & Meagher 1994). For many studies this is not a satisfactory assumption. Complete sampling of natural populations is difficult, and it is unreasonable to exclude males from paternity purely on the basis that they are unsampled. Tackling the problems posed by unsampled candidate males has been one of the major motivations for developing our simulation approach. The key assumption is that the distribution of reproductive success is the same for unsampled males as it is for sampled males. While this may not always be true, for example if alternative male mating strategies make some males easier to sample than others, any other assumption in the model would be hard to justify because appropriate data are unlikely to be available. If the proportion of candidate males sampled is low, or if there are a priori reasons for suspecting that sampled males and unsampled males have different distributions of reproductive success, caution should be exercised in extrapolating results of paternity inference using sampled individuals to the population as a whole. We also assume that the genotypes of sampled males are representative of the genotypes of unsampled males.

It is common practice in human paternity testing in some countries to use paternity likelihood ratios within a framework of Bayesian inference (Valentin 1980). Advocates of Bayesian inference point out that the approach allows prior information on paternity (the prior probability) to be combined with the genetic likelihood of paternity to derive a combined probability of paternity (the posterior probability). This leads immediately to the question: what is an appropriate prior probability? The simple answer is that nobody knows, as revealed by the furious debate that this question, and other questions on the appropriateness of Bayesian inference in paternity testing, have provoked in the literature (Walker 1983; Aickin 1984; Li & Chakravarti 1985; Elston 1986a; Elston 1986b; Li & Chakravarti 1986; Thompson 1986; Valentin 1986). We believe that our approach arrives at a statistically reasonable solution to the problem of evaluating confidence in paternity assignments without making unjustifiable assumptions about the prior probability of paternity of different males, and we therefore do not believe that a Bayesian framework is necessary or helpful.

Allowing for errors in likelihood calculations renders paternity inference relatively insensitive to typing errors. A system based on principles of exclusion may exclude the true father at relatively high frequency even when errors are infrequent, because a single typing error in any of the mother–father–offspring trio at any of the marker loci can lead to exclusion. In large-scale typing of allozymes, errors occur at a rate of the order of 1% (Lathrop *et al*. 1983), and if similar rates are true in large-scale screening of microsatellites, typing errors are likely to be a major cause of mismatches between offspring and their true parents. Another recent study (SanCristobal & Chevalet 1997) examined the effect of mistyping offspring alleles on likelihood-based paternity inference, both analytically and by simulation. It was found that allowing for errors was important, but that the choice of error rate (providing the value chosen was greater than zero) did not have a major impact on confidence or success rate, whether or not the true error rate was greater than zero. However, we believe that SanCristobal and Chevalet's error model, where individual alleles were randomly replaced with nonidentical alleles without reference to their respective frequencies, is not as realistic as our error model, where genotypes are randomly replaced by genotypes selected under Hardy–Weinberg assumptions, so that rare genotypes are more likely to be altered by a typing error than very common genotypes.

There are other possible causes of mismatches between parents and their offspring, aside from typing error. A commonly encountered problem is the presence of null alleles. At high frequencies these leave a characteristic signature of repeated homozygote–homozygote mismatches between known parent–offspring dyads, and typing of the affected locus may be discontinued (Pemberton *et al*. 1995). Null alleles at low frequencies are harder to detect, but may be treated as typing errors. Although not statistically ideal, treating a mismatch generated by a null allele as an error is preferable to treating it as a basis for exclusion. Another possible cause of mismatches between offspring and their true parents is mutation. Although mutations are alterations of single alleles rather than pairs of alleles and may not be independent of previous allelic state (e.g. the stepwise mutation model for microsatellite markers, Valdes *et al*. 1993) treating a mutation as an error is preferable to using it as a basis for exclusion.

A statistic for paternity assignment that is insensitive to the relatedness structure of the population under scrutiny is very desirable, as it is the relatedness structure that the paternity inference is designed to reveal. Simulations suggest that paternity inference using Δ is, in general, robust to the presence of unknown close relatives of the parents among the candidate males. In other words, no prior knowledge of the relatedness structure of the population is needed before Δ can be used in paternity inference. Furthermore, the concern expressed by Thompson (1976a, 1976b) and Thompson & Meagher (1987) that nonexcluded full sibs of the offspring on average have higher LOD scores than the true father does not in practice invalidate the use of Δ in analysis of paternity in most natural populations for at least three reasons. First, full sibs of the offspring (sons of the true father and the same mother) will in many populations not be considered as candidate males (e.g. because the reproductive lifespan of males is less than the time for development from conception to breeding status). Second, individuals related to one parent only do not present particular difficulties. Third, half sibs of the parents, probably the most common closest relative in polygynous species, have only modest effects on paternity inference even when present in large numbers. In conclusion, close relatives do lead to overestimation of confidence, but the overestimate is small under many commonly encountered conditions. The exception is when full sibs of the offspring (and also half sibs of the offspring when mothers are unsampled) can be considered candidate males for paternity.

The results of the simulation are useful for identifying ways to improve the success of paternity inference. For example, one can explore whether or not maternal sampling is important for successful paternity inference and whether gaps or errors in the data set are limiting factors. A preliminary screen of a small sample of the population (e.g. 25 individuals) at the chosen marker loci could also be used to predict whether or not paternity inference on the full population is likely to be productive. If the markers appear to be insufficiently informative, the simulation could be used to estimate how many additional markers would be needed to achieve a given level of success in paternity inference.

The primary purpose of this development is to aid parentage studies in natural animal and plant populations, although our method is equally applicable to captive or domestic animals where they are held in sizeable groups, and to cultivated plants, providing that populations are in Hardy–Weinberg equilibrium. The same approach to paternity inference has been applied in harbour seals (*Phoca vitulina*; Coltman *et al*., in press), and a similar approach, using an earlier version of the CERVUS program, has been used to infer paternity in Soay sheep (*Ovis aries*; Pemberton *et al*. 1996), grey seals (*Halichoerus grypus*; P. J. Allen, unpublished data) and humpback whales (*Megaptera novaeangliae*; Valsecchi 1996). Our method offers a flexible and practical framework for accurate assessment of paternity at the individual level, and we believe that it will be a useful tool in paternity inference for a wide range of species.

Acknowledgements

We thank our colleagues Tim Clutton-Brock, Steve Albon and Fiona Guinness, several field assistants and numerous volunteers who have helped to catch and sample deer over the last 25 years. We also thank Scottish Natural Heritage for permission to work on Rum, and NERC and BBSRC for financial support. Dave Coltman, Simon Goodman, Steve Paterson, Katrina Lythgoe, Nick Barton, Tom Meagher, Ruth Shaw, Terry Burke and three anonymous referees provided helpful comments on earlier drafts of this manuscript.

References

Aickin M (1984) Some fallacies in the computation of paternity probabilities. *American Journal of Human Genetics*, **36**, 904–915.

Aitkin CGG (1995) *Statistics And The Evaluation Of Evidence For Forensic Scientists*. John Wiley & Sons, Chichester, UK.

Ashton GC (1980) Mismatches in genetic markers in a large family study. *American Journal of Human Genetics*, **32**, 601–613.

Bancroft DR, Pemberton JM, King P (1995) Extensive protein and microsatellite variability in an isolated, cyclic ungulate population. *Heredity*, **74**, 326–336.

Barton NH, Gale KS (1993) Genetic analysis of hybrid zones. In: *Hybrid Zones and the Evolutionary Process*, (ed. Harrison RG), pp. 13–45 . Oxford University Press, New York.

Brenner CH (1997) Symbolic kinship program. *Genetics*, **145**, 535–542.

Buchanan FC, Crawford AM (1993) Ovine microsatellites at the OarFCB11, Oar FCB128, OarFCB193, OarFCB266 and OarFCB304 loci. *Animal Genetics*, **24**, 145.

Callen DF, Thompson AD, Shen Y, Phillips HA, Richards RI *et al.* (1993) Incidence and origin of 'null' alleles in the $(AC)_n$ microsatellite markers. *American Journal of Human Genetics*, **52**, 922–927.

Chakraborty R, Meagher TR, Smouse PE (1988) Parentage analysis with genetic markers in natural populations. I. The expected proportion of offspring with unambiguous paternity. *Genetics*, **118**, 527–536.

Chakraborty R, Shaw M, Schull WJ (1974) Exclusion of paternity: the current state of the art. *American Journal of Human Genetics*, **26**, 477–488.

Chakravarti A, Li CC (1983) The effect of linkage on paternity calculations. In: *Inclusion Probabilities In Parentage Testing*, (ed. Walker RH), pp. 411–422. American Association of Blood Banks, Arlington, Virginia.

Clutton-Brock TH, Albon SD (1989) *Red Deer In The Highlands*. BSB Professional Books, Oxford.

Clutton-Brock TH, Guinness FE, Albon SD (1982) *Red Deer – Behavior And Ecology Of Two Sexes*. Edinburgh University Press, Edinburgh.

Coltman DW, Bowen WD, Wright JM (1998) Male mating success in an aquatically mating pinniped, the harbour seal (*Phoca vitulina*) determined by microsatellite DNA markers. *Molecular Ecology*, **7**, 627–638.

Devlin B, Roeder K, Ellstrand NC (1988) Fractional paternity assignment: theoretical development and comparison to other methods. *Theoretical and Applied Genetics*, **76**, 369–380.

Ede AJ, Pierson CA, Crawford AM (1995) Ovine microsatellites at the OarCP9, OarCP16, OarCP20, OarCP21, OarCP23 and OarCP26 loci. *Animal Genetics*, **26**, 129–130.

Edwards AWF (1972) *Likelihood*. Cambridge University Press, Cambridge.

Elston RC (1986a) Probability and paternity testing. *American Journal of Human Genetics*, **39**, 112–122.

Elston RC (1986b) Some fallacious thinking about the paternity index: comments. *American Journal of Human Genetics*, **39**, 670–672.

Foltz DW, Hoogland DW (1981) Analysis of the mating system in the black-tailed prairie dog (*Cynomys ludovicianus*) by likelihood of paternity. *Journal of Mammalogy*, **62**, 706–712.

Georges M, Massey J (1992) *Polymorphic DNA Markers In Bovidae*. World Intellectual Property Org., Geneva.

Hogg JT, Forbes SH (1997) Mating in bighorn sheep: frequent male reproduction via a high-risk 'unconventional' tactic. *Behavioral Ecology and Sociobiology*, **41**, 33–48.

Jamieson A (1965) The genetics of transferrins in cattle. *Heredity*, **20**, 419–441.

Jamieson A (1994) The effectiveness of using co-dominant polymorphic allelic series for (1) checking pedigrees and (2) distinguishing full-sib pair members. *Animal Genetics*, **25**, 37–44.

Jones AG, Avise JC (1997) Microsatellite analysis of maternity and the mating system in the Gulf pipefish *Syngnathus scovelli*, a species with male pregnancy and sex-role reversal. *Molecular Ecology*, **6**, 203–213.

Keane B, Dittus WPJ, Melnick DJ (1997) Paternity assignment in wild groups of toque macaques *Macaca sinica* at Polonnaruwa, Sri Lanka using molecular markers. *Molecular Ecology*, **6**, 267–282.

Lathrop GM, Hooper AB, Huntsman JW, Ward RH (1983) Evaluating pedigree data. I. The estimation of pedigree error in the presence of marker mistyping. *American Journal of Human Genetics*, **35**, 241–262.

Li CC, Chakravarti A (1985) Basic fallacies in the formulation of the paternity index. *American Journal of Human Genetics*, **37**, 809–818.

Li CC, Chakravarti A (1986) Some fallacious thinking about the paternity index: a reply to Dr Jack Valentin's comments. *American Journal of Human Genetics*, **38**, 586–589.

Meagher TR (1986) Analysis of paternity within a natural population of *Chamaelirium luteum*. I. Identification of most-likely male parents. *The American Naturalist*, **128**, 199–215.

Morin PA, Wallis J, Moore JJ, Woodruff DS (1994) Paternity exclusion in a community of wild chimpanzees using hypervariable simple sequence repeats. *Molecular Ecology*, **3**, 469–478.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Pemberton JM, Albon SD, Guinness FE, Clutton-Brock TH, Berry RJ (1988) Genetic variation and juvenile survival in red deer. *Evolution*, **42**, 921–934.

Pemberton JM, Albon SD, Guinness FE, Clutton-Brock TH, Dover GA (1992) Behavioral estimates of male mating success tested by DNA fingerprinting in a polygynous mammal. *Behavioral Ecology*, **3**, 66–75.

Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Non-amplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*, **4**, 249–252.

Pemberton JM, Smith JA, Coulson TN, Marshall TC, Slate J *et al.* (1996) The maintenance of genetic polymorphism in small island populations: large mammals in the Hebrides. *Philosophical Transactions of the Royal Society London Series B*, **351**, 745–752.

Pena SDJ, Chakraborty R (1994) Paternity testing in the DNA era. *Trends in Genetics*, **10**, 204–209.

Phillips HA, Thompson AD, Kozman HM, Sutherland GR, Mulley JC (1993) A microsatellite marker within the duplicated D16S79 locus has a null allele: significance for linkage mapping. *Cytogenetics and Cell Genetics*, **64**, 131–132.

Queller DC, Strassmann JE, Hughes CR (1993) Microsatellites and kinship. *Trends in Ecology and Evolution*, **8**, 285–288.

Roeder K, Devlin B, Lindsay BG (1989) Application of maximum likelihood methods to population genetic data for the estimation of individual fertilities. *Biometrics*, **45**, 363–379.

SanCristobal M, Chevalet C (1997) Error tolerant parent identification from a finite set of individuals. *Genetical Research*, **70**, 53–62.

Smouse PE, Meagher TR (1994) Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L.) Gray (Liliaceae). *Genetics*, **136**, 313–322.

Swarbrick PA, Buchanan FC, Crawford AM (1991) Ovine dinucleotide repeat polymorphism at the MAF35 locus. *Animal Genetics*, **22**, 369–370.

Swarbrick PA, Crawford AM (1992) Ovine dinucleotide repeat polymorphism at the MAF109 locus. *Animal Genetics*, **23**, 84.

Taylor AC, Horsup A, Johnson CN, Sunnucks P, Sherwin B (1997) Relatedness structure detected by microsatellite analysis and attempted pedigree reconstruction in an endangered marsupial, the northern hairy-nosed wombat *Lasiorhinus krefftii*. *Molecular Ecology*, **6**, 9–19.

Thompson EA (1975) The estimation of pairwise relationships. *Annals of Human Genetics*, **39**, 173–188.

Thompson EA (1976a) Inference of genealogical structure. *Social Science Information*, **15**, 477–526.

Thompson EA (1976b) A paradox of genealogical inference. *Advances in Applied Probability*, **8**, 648–650.

Thompson EA (1986) Likelihood inference of paternity. *American Journal of Human Genetics*, **39**, 285–287.

Thompson EA, Meagher TR (1987) Parental and sib likelihoods in genealogy reconstruction. *Biometrics*, **43**, 585–600.

Valdes AM, Slatkin M, Freimer NB (1993) Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics*, **133**, 737–749.

Valentin J (1980) Exclusions and attributions of paternity: practical experiences of forensic genetics and statistics. *American Journal of Human Genetics*, **32**, 420–431.

Valentin J (1986) Some fallacious thinking about the paternity index. *American Journal of Human Genetics*, **38**, 582–585.

Valsecchi E (1996) *Genetic Analysis Of The Humpback Whale (Megaptera novaeangliae) Using Microsatellites*. PhD thesis, University of Cambridge, UK.

Walker RH (1983) *Inclusion Probabilities In Parentage Testing*. American Association of Blood Banks, Arlington, Virginia.

The authors are involved in various analyses, several depending on paternity inference, using the individually monitored red deer population on Rum and Soay sheep population on St Kilda. Tristan Marshall is a BBSRC-funded PhD student who is investigating inbreeding and its consequences in Rum red deer and reintroduced Arabian oryx in Oman. As part of his PhD work, he has developed the techniques for paternity inference described in this paper and written the CERVUS program.

Appendix 1

Derivation of likelihoods that take into account typing errors

Defining x_i to be the unknown, true genotype of individual i that has been mistyped at the locus under consideration, eqn 2 becomes:

$$L(H_1|g_m,g_a,g_o) = (1-e)^3.T(g_o|g_m,g_a).P(g_m).P(g_a) +$$
$$e(1-e)^2 \begin{bmatrix} T(g_o|g_m,x_a).P(g_m).P(x_a) + \\ T(g_o|x_m,g_a).P(x_m).P(g_a) + \\ T(x_o|g_m,g_a).P(g_m).P(g_a) \end{bmatrix} +$$
$$e^2(1-e) \begin{bmatrix} T(x_o|g_m,x_a).P(g_m).P(x_a) + \\ T(x_o|x_m,g_a).P(x_m).P(g_a) + \\ T(g_o|x_m,x_a).P(x_m).P(x_a) \end{bmatrix} +$$
$$e^3.T(x_o|x_m,x_a).P(x_m).P(x_a)$$

where e is the error rate. As x_i can be any genotype, $P(x_i) = 1$ and $T(g_j|x_i) = P(g_j)$ for any i, j, and so this simplifies to:

$$L(H_1|g_m,g_a,g_o) = (1-e)^3.T(g_o|g_m,g_a).P(g_m).P(g_a) + \quad (6)$$
$$e(1-e)^2 \begin{bmatrix} T(g_o|g_m).P(g_m) + \\ T(g_o|g_a).P(g_a) + P(g_m).P(g_a) \end{bmatrix} +$$
$$e^2(1-e) \quad [P(g_m) + P(g_a) + P(g_o)] + e^3$$

Eqn 3 becomes:

$$L(H_2|g_m,g_a,g_o) = (1-e)^3.T(g_o|g_m).P(g_m).P(g_a) +$$
$$e(1-e)^2 \begin{bmatrix} T(g_o|g_m).P(g_m).P(x_a) + \\ T(g_o|x_m).P(x_m).P(g_a) + \\ T(x_o|g_m).P(g_m).P(g_a) \end{bmatrix} +$$
$$e^2(1-e) \begin{bmatrix} T(x_o|g_m).P(g_m).P(x_a) + \\ T(x_o|x_m).P(x_m).P(g_a) + \\ T(g_o|x_m).P(x_m).P(x_a) \end{bmatrix} +$$
$$e^3.T(x_o|x_m).P(x_m).P(x_a)$$

which simplifies to:

$$L(H_2|g_m,g_a,g_o) = (1-e)^3.T(g_o|g_m).P(g_m).P(g_a) + \quad (7)$$
$$e(1-e)^2 \begin{bmatrix} T(g_o|g_m).P(g_m) + \\ P(g_o).P(g_a) + P(g_m).P(g_a) \end{bmatrix} +$$
$$e^2(1-e) \quad [P(g_m) + P(g_a) + P(g_o)] + e^3$$

The likelihood ratio is eqn 6 divided by eqn 7, as before, and reduces to eqn 4 when $e = 0$.

When mothers are unsampled, errors may occur in the offspring or the alleged father. The likelihood that the alleged father is the true father is:

$$L(H_1|g_a,g_o) = (1-e)^2.T(g_o|g_a).P(g_a) +$$
$$e(1-e)^2\,[\,T(g_o|x_a).P(x_a) + T(x_o|g_a).P(g_a)] +$$
$$e^2.T(x_o|x_a).P(x_a)$$

which reduces to:

$$L(H_1|g_a,g_o) = (1-e)^2.T(g_o|g_a).P(g_a) + e(1-e)[P(g_o) + P(g_a)] + e^2 \quad (8)$$

The likelihood that the alleged father is a male selected at random is:

$$L(H_2|g_a,g_o) = (1-e)^2.P(g_o).P(g_a) +$$
$$e(1-e)\,[\,P(g_o).P(x_a) + P(x_o).P(g_a)] +$$
$$e^2.P(x_o).P(x_a)$$

which is:

$$L(H_2|g_a,g_o) = (1-e)^2.P(g_o).P(g_a) + e(1-e)[P(g_o) + P(g_a)] + e^2 \quad (9)$$

The likelihood ratio is eqn 8 divided by eqn 9, and reduces to eqn 5 when $e = 0$.

Appendix 2

Average probability of exclusion

The average probability of excluding an unrelated individual from parentage, given the genotypes of the offspring and other parent and assuming Hardy–Weinberg equilibrium, was derived initially by Jamieson (1965), but is most easily calculated in the form given by Chakravarti & Li (1983) and Jamieson (1994). In Chakravarti and Li's notation, the average exclusion probability P_l at a locus l with k codominant alleles is given by:

$$P_l = a_1 - 2a_2 + a_3 + 3(a_2a_3 - a_5) - 2({a_2}^2 - a_4)$$

where

$$a_n = \sum_{i=1}^{k} {p_i}^n$$

and p_i is the frequency of allele i, and $a_1 = 1$. Note that this equation was incorrectly cited in Morin et al. (1994). The equivalent average probability of excluding an unrelated individual from parentage given only the genotype of the offspring is derived below.

For homozygous offspring AA, an exclusion occurs if the candidate parent is neither AA nor any of the $k-1$ heterozygotes AX. For heterozygous offspring AB, an exclusion occurs if the candidate parent is neither AA, BB, any of the $k-1$ heterozygotes AX nor any of the $k-1$ heterozygotes BX. The heterozygous candidate parent AB occurs both in the set of genotypes AX and the set of genotypes BX. Defining the probability of genotypes AA, AB, AX and BX as $p(ii)$, $p(ij)$, $p(ix)$ and $p(jx)$ respectively and summing across all pairwise genotypic combinations, the average probability of exclusion at locus l, P_l, can be written:

$$P_l = 1 - \left\{ \sum_{i=1}^{k} \sum_{x=1}^{k} p(ii)p(ix) + \frac{1}{2} \sum_{i \neq j}^{k} \sum_{x=1}^{k} p(ij)(p(ix) + p(jx) - p(ij)) \right\}$$

Substituting the expected frequencies of the relevant genotypes, assuming Hardy–Weinberg equilibrium, yields:

$$P_l = 1 - \left\{ \sum_{i=1}^{k} {p_i}^2 \left(\sum_{x=1}^{k} 2p_i p_x - {p_i}^2 \right) + \right.$$

$$\left. \sum_{i \neq j}^{k} p_i p_j \left(\sum_{x=1}^{k} 2p_i p_x - {p_i}^2 + \sum_{x=1}^{k} 2p_j p_x - {p_j}^2 - 2p_i p_j \right) \right\}$$

where p_i is the frequency of allele i as above. Given that

$$\sum_{i=1}^{k} p_i = 1,$$

this simplifies to:

$$P_l = 1 - \left\{ \sum_{i=1}^{k} {p_i}^2 . p_i(2 - p_i) + \sum_{i \neq j}^{k} p_i p_j (p_i + p_j)(2 - p_i - p_j) \right\}$$

By writing

$$\sum_{i \neq j}^{k} p_i p_j \text{ as } \sum_{i=1}^{k} p_i \left(\left(\sum_{j=1}^{k} p_j \right) - p_i \right)$$

and similarly for other powers, P_l becomes:

$$P_l = a_1 - 4a_2 + 4a_3 - 3a_4 + 2{a_2}^2 \qquad (10)$$

where

$$a_n = \sum_{i=1}^{k} {p_i}^n$$

as above.

The overall average probability of exclusion across n independently inherited loci, P, may be calculated in the usual way:

$$P = 1 - \prod_{l=1}^{n} [1 - P_l]$$

Estimating the rate of typing error

Equation 10 may be used in estimating the rate of typing error if a large number of parent–offspring pairs are known without error. Defining an error as the replacement of the true genotype with a genotype selected at random under Hardy–Weinberg assumptions, the rate of typing error e_l for locus l is approximately:

$$e_l \approx \frac{1}{2P_l} \cdot \frac{m_l}{M_l}$$

where m_l is the observed number of parent–offspring mismatches in M_l comparisons, assuming that the probability of both parent and offspring being mistyped is negligible. All parent–offspring pairs used in estimating e_l must be independent. One must avoid including multiple representatives of a half sibship and also avoid including an individual both as an offspring and as a parent, otherwise the error rate is liable to be overestimated.

If error is constant across loci, a better estimate of the underlying error rate, e, is the average across n loci:

$$e = \frac{1}{n} \sum_{l=1}^{n} e_l.$$