Statistical Approaches to Paternity Analysis in Natural Populations and Applications to the North Atlantic Humpback Whale

Rasmus Nielsen,* David K. Mattila,† Philip J. Clapham‡ and Per J. Palsbøll§

*Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138 and Department of Biometrics, Cornell University, Ithaca, New York 14853-7801, †Center for Coastal Studies, Provincetown, Massachusetts 02567, ‡Northeast Fisheries Science Center, Woods Hole, Massachusetts 02543 and §School of Biological Sciences, University of Wales, Gwynedd LL57 2UW, United Kingdom

ABSTRACT

We present a new method for paternity analysis in natural populations that is based on genotypic data that can take the sampling fraction of putative parents into account. The method allows paternity assignment to be performed in a decision theoretic framework. Simulations are performed to evaluate the utility and robustness of the method and to assess how many loci are necessary for reliable paternity inference. In addition we present a method for testing hypotheses regarding relative reproductive success of different ecologically or behaviorally defined groups as well as a new method for estimating the current population size of males from genotypic data. This method is an extension of the fractional paternity method to the case where only a proportion of all putative fathers have been sampled. It can also be applied to provide abundance estimates of the number of breeding males from genetic data. Throughout, the methods were applied to genotypic data collected from North Atlantic humpback whales (*Megaptera novaeangliae*) to test if the males that appear dominant during the mating season have a higher reproductive success than the subdominant males.

THE use of genetic markers to identify parent-offspring relationships is becoming an important tool in molecular ecology. In some studies the issue of paternity is of interest in itself (*e.g.*, FOLTZ and HOGLAND 1981; CLAPHAM and PALSBØLL 1997). In other cases paternity analysis is used in the estimation or detection of gene flow between populations (*e.g.*, AMOS *et al.* 1993) or the analysis of reproductive success of different ecological or behavioral groups (*e.g.*, SMOUSE and MEAGHER 1994).

The basic statistical methodology is based on the calculation of likelihoods in genealogies (THOMPSON 1975, 1976). The probability of an observed offspring genotype can be calculated knowing the parental genotypes, usually assuming Mendelian segregation of alleles. Calculation of this probability for multiple potential fathers provides the likelihood function for a single offspring, and paternity can be assigned by choosing the most likely father among the potential fathers. This type of approach has been developed and applied by MEAGHER (1986) and MEAGHER and THOMPSON (1986, 1987). One of the key questions relating to these methods is how to assess the confidence of a particular paternity assignment. In the (now) commonly applied approach developed by MARSHALL *et al.* (1998) the likelihood

values of the two potential fathers with the highest likelihood values are compared and the logarithm of the ratio of these two likelihood values is treated as a test statistic (Δ). The significance of the difference in likelihood estimates is assessed by estimating the null-distribution of Δ from simulations. If the observed value of Δ is sufficiently large, the potential father with the highest likelihood is accepted as the father. This approach was developed as a method for assigning paternity when more than one male cannot be excluded by the data.

The likelihood approach by MARSHALL *et al.* (1998) may be improved upon for several reasons. First, Δ may not be the best statistic for assigning paternity, since it ignores information regarding all potential fathers apart from the two with the highest likelihood values. Also, in many cases, it may not be of interest to make a binary decision regarding parentage. Often the relevant biological question is to assess the relative reproductive success of different geographically, ecologically, or behaviorally defined groups. For this purpose, methods known as fractional assignment methods have been developed (DEVLIN *et al.* 1988; ROEDER *et al.* 1989; SMOUSE and MEAGHER 1994). In these approaches, reproductive success is estimated by weighting the reproductive contribution of a potential parent with the likelihood of paternity of the parent. As mentioned by ROEDER *et al.* (1989), this approach can be considered a Bayesian procedure in which all parents are given equal prior probability of paternity. One of the advantages of the fractional-likelihood approach is that the likelihood

Corresponding author: Rasmus Nielsen, Department of Biometrics, 439 Warren Hall, Cornell University, Ithaca, NY 14853-7801. E-mail: rn28@cornell.edu

function for a specific parameter relating to the reproductive success or dispersal of different groups can be calculated directly from the data.

In this article we present an approach for estimating parentage probabilities, which can be considered a Bayesian alternative to the method developed by MAR-SHALL *et al.* (1998). The method proceeds by making inferences directly on the basis of the calculated parentage probabilities. We here use the term "parentage probability" to describe the posterior probability that a particular putative father is the actual father. Subsequently, we develop a method for testing hypotheses regarding reproductive success and for estimating population sizes on the basis of parent-offspring genotypic data. This method can be viewed as an extension of the aforementioned fractional paternity approach to the case where only a proportion of all potential males have been sampled. Previous approaches implicitly assume that all individuals in the population have been sampled. We show that inferences regarding paternity are highly sensitive to the sampling fraction but may be surprisingly robust to violations of the underlying assumptions regarding family structure.

The method developed here is applied to genotypic data obtained from North Atlantic humpback whales, *Megaptera novaeangliae*. In the case of cetaceans (whales, dolphins, and porpoises) maternity is readily inferred from the close association between the mother and her calf before the calf is weaned, whereas paternity is almost impossible to infer from observation alone. Thus for paternity assessment, genetic analyses appear to be the only viable method to evaluate reproductive success, but only a handful of studies have employed genotypic data toward this objective in cetaceans so far (AMOS *et al.* 1991, 1993; CLAPHAM and PALSBØLL 1997). The issue of mating behavior and male reproductive success is particularly difficult to assess in the baleen whales, which do not exhibit the tight and well-defined pod structure often observed among toothed whales. In addition, only a few behaviors among baleen whales can be directly or indirectly related to mating (and these only in a few species).

POSTERIOR PROBABILITIES OF PATERNITY

Our objective is to estimate the posterior probability that a particular individual might be the father of a known offspring. We use the posterior probability of paternity directly to measure our belief in the paternity assessment. In this sense, the method can be viewed as a Bayesian method for paternity inference.

We assume multiple mother-offspring pairs as well as multiple potential fathers and we allow for the possibility that not all potential fathers in the breeding population have been sampled. To estimate the probability that a potential father is the father of an offspring we need to make assumptions regarding the prior probability of

a potential father being the father. In the absence of other information, we assume that the prior probability that a particular male is the father is $1/N$, where N is the number of potentially breeding males in the breeding area. We note that in some circumstances this may not be the best prior to use. In some cases there might be other information available, for example, regarding population subdivision or age structure, which might suggest that not all males in the population have the same probability of siring an offspring. The method we describe can easily be adjusted in such situations to take this information into account.

Let $I_j(i)$ indicate the event that the jth potential father is the father of the ith offspring. Also, let the ith maternal genotype be $\mathbf{M}_i$, the associated genotype of the offspring be $\mathbf{O}_i$, the genotype of jth potential father be $\mathbf{F}_j$, and $\mathbf{A}$ be the matrix of allelic frequencies for all loci. If we have sampled n of N males on the breeding ground (N is assumed to be large), the posterior probability of paternity can be calculated as

$$\Pr(I_k(i) \mid \mathbf{M}_i, \mathbf{F}, \mathbf{A}, N) = \frac{\Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}_k)}{\sum_{j=1}^{n}\Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}_j) + (N-n)\Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{A})},$$
(1)

where $\Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}_j)$ is the shorthand notation for $\Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}_j, I_j(i))$. Assuming Mendelian segregation and independence among loci we can easily calculate the probability of an observed offspring genotype given the maternal genotype and the genotype of a particular potential father $\Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}_j)$, using standard methods (*e.g.*, THOMPSON 1975, 1976). Likewise, $\Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{A})$ can easily be calculated assuming Hardy-Weinberg equilibrium and independence among loci (linkage equilibrium). To perform this calculation, the population allele frequencies ($\mathbf{A}$) must be known. Although these frequencies will rarely or never be known in natural populations, estimates of the observed allelic frequencies can be used in place of the population frequencies for large samples. This method also requires information regarding the number of breeding males in the population. In some cases such information is available through direct estimates of population census size. Cases where such information is not available are also treated.

PATERNITY INFERENCE WHEN THE POPULATION SIZE IS UNKNOWN

In many cases, the problem of identifying parent-offspring relationships has been presented as a problem of classifying parent-offspring relations as either a match or not a match. A given offspring can either be assigned to a sampled potential father or classified as having no father among the sampled males. Some authors have chosen to phrase the problem of confidence in a paternity assignment in terms of hypothesis testing (*e.g.*, MAR-SHALL *et al.* 1998). We instead suggest the use of an

explicit decision-theoretic approach to the problem of paternity assignment, *i.e.*, we define a specific loss function, which provides the "loss" incurred if a wrong classification is made. By minimizing the expectation of the loss (the risk), we can establish an appropriate decision rule that determines the classification of parent-offspring relationships. Using a 0–1 loss function, *i.e.*, a loss of 1 if an incorrect classification is made, the risk is simply the probability of misclassification. The posterior risk is minimized by accepting a match only if it has the largest posterior probability of any match and if the posterior probability of the match is larger than the posterior probability that the father was not sampled. This is the decision rule we use in the following. We could have chosen another loss function; for example, it might be reasonable to assign a larger loss to a misclassification in which a match between two unrelated individuals is accepted than to a misclassification in which we fail to identify a parent-offspring relationship. For example, such a loss function could lead to a decision rule in which a match is accepted if the posterior probability of paternity is >95% or 99% analogous to the criteria usually used in hypothesis testing. However, in the absence of other information regarding the application of the method, we assign the same loss function to all misclassifications.

We also note that in many biological studies it is more relevant to use the probabilities of paternity directly instead of making binary decisions regarding paternity.

PERFORMANCE ASSESSMENT—HOW MANY LOCI ARE NECESSARY?

We employed computer simulations to evaluate our approach. In these simulations, we focused on the methods' ability to make binary decisions regarding paternity as described above. Multiple data sets were generated and for each data set the proportion of correctly classified offspring-parent relationships was scored. In the simulation of a data set, the population frequencies of alleles were first determined for each of k loci. Two different sets of population frequencies were used for the simulations: 10 alleles each of frequency 0.1 and 4 alleles each of frequency 0.25. Subsequently, a set of c maternal genotypes was generated and a set of n male genotypes was generated independently. Offspring genotypes were generated by randomly choosing one allele for each locus from the mother and with probability n/N choosing paternal alleles from a father among the n male genotypes, and with probability $(N - n)/N$ choosing the paternal alleles by sampling from the population frequencies. Throughout the simulations we assumed Mendelian segregation and independence among loci.

For each generated data set, the population allele frequencies were estimated from the observed allele frequencies. Paternity analysis was then performed as
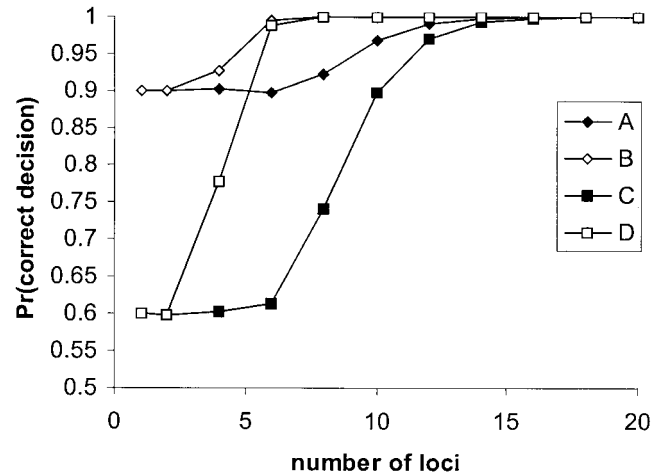

FIGURE 1.—Proportion of correct paternity decisions. Parameter values: (A) 4 alleles with equal frequencies, $c = 50$ and $n = 50$; (B) 10 alleles with uniform frequencies, $c = 50$ and $n = 50$; (C) 4 alleles with equal frequencies, $c = 200$ and $n = 200$; (D) 10 alleles with equal frequencies, $c = 200$ and $n = 200$. In all cases $N = 500$ and 1000 simulations were performed.

described above. A total of 1000 simulations were performed for each parameter value and the proportion of offspring that were correctly classified was scored as a measure of the performance of the method. It was assumed that $N = 500$ and n, k, and c were varied to examine the performance of the method under multiple parameter settings.

The results of the simulations are presented in Figure 1. Our results differed considerably between the two levels of variation. In the case of 10 alleles as few as six loci are sufficient for reliable paternity inference given the sample and population size employed in our simulations. In contrast, in the case of 4 alleles, as many as 10–14 loci are necessary for reliable paternity inference. Clearly, the variability of the locus is a major determining factor of the performance of the method.

The allelic frequencies used in the simulations are idealized. The equal distribution of allele frequencies would, for example, be expected under a k-allele model with symmetric mutation and very high mutation rates. However, in real data (*e.g.*, microsatellite data) the allele frequencies are unlikely to follow such a distribution and thus considerably more alleles are required at each locus to yield an equal performance.

ROBUSTNESS

The method described above is an improvement of previous methods in that it takes incomplete sampling of putative fathers into account. However, it shares some of the problems of previous methods in making very simple assumptions regarding family structure. Most importantly, it ignores the possibility that some of the
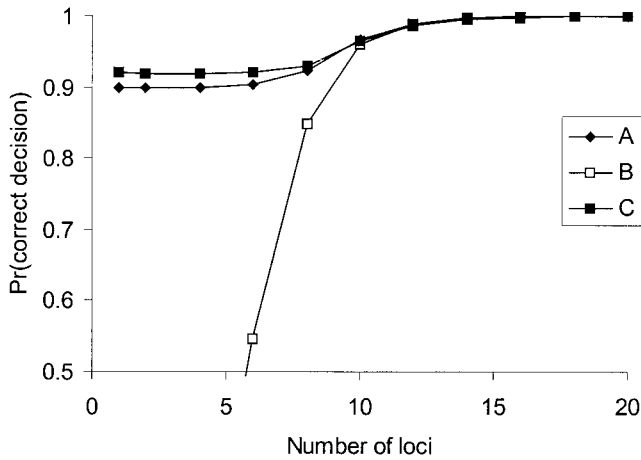
FIGURE 2.—Proportion of correct paternity decisions. The proportion of all paternity decisions (offspring assigned to a putative male or assigned to nonsampled males) that are correct is shown. The data were simulated assuming $c = 50$, $n = 50$, and $N = 500$. In A and B none of the putative fathers are half-sibs to individuals in the offspring generation. In C 20% of all putative fathers are paternal half-sibs of individuals in the offspring generation and 25% of all males have all the offspring (75% of all males have fertility 0). In B it was incorrectly assumed that $N = n$, *i.e.*, unobserved males are not taken into account. In all cases $N = 500$ and 1000 replicate data sets were simulated.

potential fathers may actually be siblings or other relatives of the sampled offspring. Also, it relies on the assumption of equal fertilities (potential for reproductive success). Here we are interested in assessing how important the problem of ignoring family structure and variation in fertility is *vs.* the importance of ignoring incomplete sampling. We do this by performing computer simulations that include incomplete sampling and family structure and determine how well the method for parentage assignment performs. We model the problem of family structure by including a proportion of paternal sibs as putative fathers. In humpback whales (in which we apply our methods later) matings have been shown to be promiscuous and full sibs are probably rare (CLAPHAM and PALSBØLL 1997). Hence, varying the proportion of paternal half sibs in the offspring data seems to be an appropriate way to examine the effect of family structure. The simulations were performed by first generating a set of $c + c_f$ maternal genotypes and a set N paternal genotypes. Offspring genotypes were then generated by choosing a random mate among the N males for each of the first c maternal genotypes. The probability that each particular male fathered an offspring was given by the relative fertility of the male. For example, in the case of equal fertility, each of the males had probability $1/N$ of fathering a particular offspring. The first c generated maternal genotypes and offspring genotypes were included as maternal and offspring data. A set of c_f offspring was similarly generated and included in the sample among the n potential fathers. The fraction of half sibs generated in this way among the n
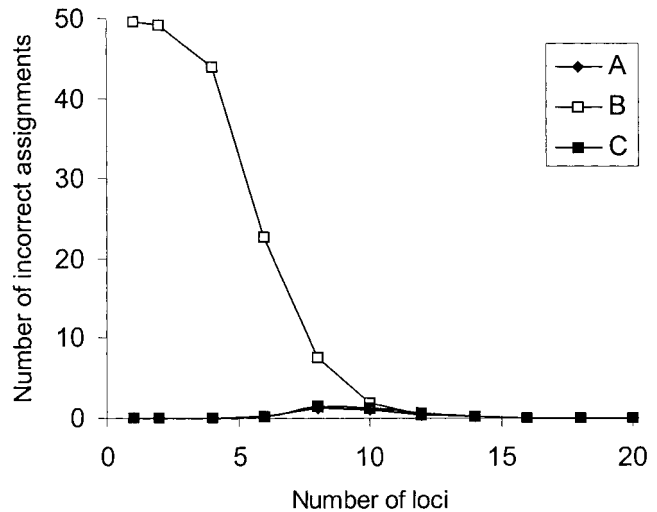

FIGURE 3.—Incorrect paternity assignments in the presence of family structure. The average number of incorrect paternity assignments in the presence of family structure using a 0–1 loss function is shown. The data were simulated under the same conditions as in Figure 2.

potential fathers is $f = c_f/n$, $0 \le f \le 1$. In this manner it is possible to examine the effect of family structure in terms of half sibs among the potential fathers and the effect of unequal reproductive success in addition to the effect of population size.

We first consider the case of equal reproductive success and no half sibs among the n male genotypes ($f = 0$). Figure 2A shows the proportion of correct paternity decisions when $c = 50$, $n = 50$, $N = 500$ and there are four alleles in each locus. Note the similarity to Figure 1A. Next (Figure 2B), we assigned paternities by wrongly assuming that the sample size equals the number of breeding males ($n = N$). As expected, the probability of a correct decision is dramatically decreased. We can also examine the number of matches incorrectly inferred (Figure 3, A and B). As expected, we see that ignoring the presence of unobserved males gives too many false matches. The effect can be very drastic even for moderate amounts of genetic data. For example, for 6 loci the average number of incorrectly inferred matches is increased >100-fold. Even for 10 loci the number of incorrectly inferred matches is almost doubled. In other words, ignoring unobserved males, as has been common in some previous methods, has a very strong effect on the number of incorrect assignments, even with moderate amounts of genetic data.

Next, we examine the effect of ignoring the possibility of unequal fertilities (wrong prior) and of family structure. We do this by letting 20% of the sampled potential fathers be paternal sibs to individuals in the offspring generation and by letting one-quarter of males sire all the offspring, *i.e.*, three-quarters of all males in the population sire none at all (Figures 2C and 3C). These violations of the assumptions of the method lead to a very small decrease in the number of correctly matched indi-
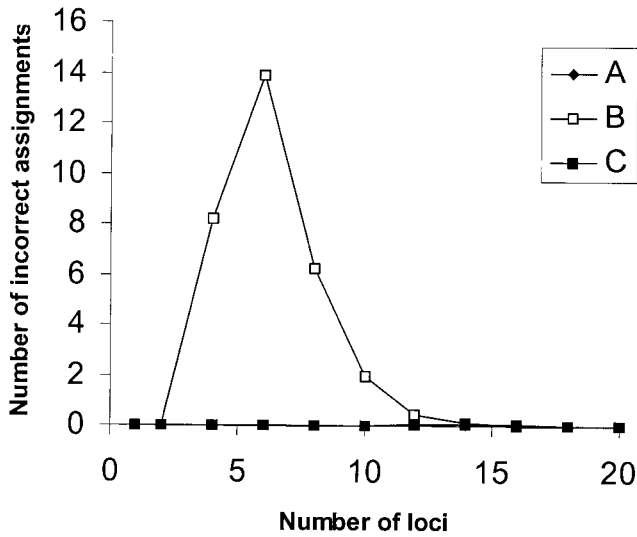
FIGURE 4.—Incorrect assignments with a 95% decision rule. The average number of incorrect paternity assignments when a match is assigned if the posterior probability of paternity is >95% is shown. The data were simulated under the same conditions as in Figures 2 and 3.

viduals and an increase in the number of incorrectly matched individuals. However, since there are fewer total father-offspring pairs, the probability of a correct decision is increased.

The effect of ignoring unobserved males on the number of incorrect matches is orders of magnitude larger than the effect of family structure. The most critical model assumptions are obviously the assumptions regarding complete sampling and the number of breeding males.

To show that this conclusion is not just a result of the chosen decision rule, we also performed simulations using another decision rule. In these simulations a match was assigned if the posterior probability of paternity was >95%. The number of incorrectly inferred paternities is shown in Figure 4. Note again that there is a drastic reduction in the performance of the method when the presence of unobserved males is ignored. In contrast, the effect of family structure and variance in the fertility among males is negligible. These conclusions cannot be guaranteed to hold for all types of paternity inference. For example, in some applications in forensic science, the presence of family structure may be of strong importance. However, for the purpose of paternity inference and assessment of fertilities in the present framework, it seems safe to conclude that family structure and variance in the fertility among males is a very minor problem compared to the problem regarding unobserved males.

TESTING HYPOTHESES REGARDING REPRODUCTIVE SUCCESS

In many cases, it is of interest to test hypotheses regarding the relative reproductive success of different ecologically or behaviorally defined groups. For example, let us assume that there are two groups, group 1 and group 2, and that we are interested in testing if the reproductive success of the two groups differs. Assume that the ratio of the reproductive success of groups 2 and 1 is α. If both groups have the same reproductive success α will equal 1 whereas, for instance, $\alpha = 3$ implies that the reproductive success of group 2 is three times larger than that of group 1. Our aim is then to obtain an estimate of α and to test the null hypothesis of $\alpha = 1$. We use a likelihood approach similar to that presented by SMOUSE and MEAGHER (1994); however, the method is modified to account for the fact that not all the potential fathers have been sampled. It is a natural extension of the Bayesian approach for classifying parent-offspring relationships described above. Assuming that the probabilities of an individual male siring two offspring are independent, the likelihood function for α is given by

$$L(\alpha) \propto \prod_{i=1}^{k} \Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \alpha) \qquad (2)$$

in a sample containing k offspring. $\mathbf{F}^{(s)}$ is the vector of genotypes of potential fathers belonging to group s, $s = 1, 2$. Let us denote the event that the father of the ith offspring is sampled and belongs to group i in the sample by $I_s(i)$, $s = 1, 2$, and the event that the father is not in the sample by $I_0(i)$. Then

$$\Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \alpha) = \Pr[\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}^{(1)}, I_1(i)]\Pr[I_1(i) \mid \alpha]$$
$$+ \Pr[\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}^{(2)}, I_2(i)]\Pr[I_2(i) \mid \alpha]$$
$$+ \Pr[\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{A}]\Pr[I_0(i)]. \qquad (3)$$

Assuming that the probability of obtaining the father in the sample equals the sampling fraction, we have

$$\Pr(I_0(i)) = (N - n_1 - n_2)/N$$
$$\Pr(I_1(i) \mid \alpha) = \alpha n_1(1 - \Pr(I_0(i)))/(\alpha n_1 + n_2)$$
$$\Pr(I_2(i) \mid \alpha) = n_2(1 - \Pr(I_0(i)))/(\alpha n_1 + n_2). \qquad (4)$$

$\Pr[\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}^{(1)}, I_1(i)]$ and $\Pr[\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}^{(2)}, I_2(i)]$ can easily be calculated as

$$\Pr[\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}^{(s)}, I_s(i)] = \frac{1}{n_s} \sum_{j=1}^{n_s} \Pr[\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}_j^{(s)}], \quad s = 1, 2, \qquad (5)$$

where $\mathbf{F}_j^{(s)}$ is the genotype of the jth potential father in group s. $\Pr[\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}_j^{(s)}]$, $s = 1, 2$, and $\Pr[\mathbf{O}_i \mid \mathbf{A}]$ can be estimated as before. Using this method, we can estimate α and perform hypothesis tests using a standard likelihood-ratio test. Numerical optimization of the likelihood function is easily done using standard methods, in this case a quasi-Newton method (PRESS et al. 1988, pp. 425ff).

In most cases, special interest focuses on testing the hypothesis of equal fertilities ($\alpha = 1$). To perform this likelihood-ratio test, some care must be taken. We note that as the number of loci grows large the likelihood function will converge to a multinomial distribution
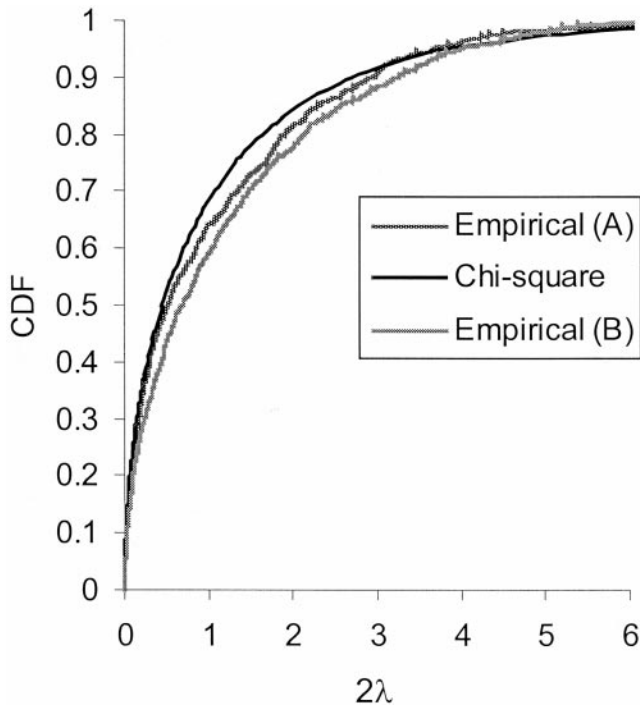
FIGURE 5.—Fit of the χ_1^2 approximation. The empirical cumulative distribution function (CDF) of the likelihood-ratio test statistic under the null-hypothesis from simulated data and the CDF of a χ_1^2 distribution, when (A) no family structure is assumed and (B) it is assumed that 20% of all putative fathers are half-sibs to the offspring generation. The hypothesis being tested is $\alpha = 1.0$. A total of 1000 simulations were used to generate each of the empirical CDFs.

with parameters $\Pr(I_0(i))$, $\Pr(I_1(i)|\alpha)$, and $\Pr(I_2(i)|\alpha)$. The standard limiting results for the likelihood function should therefore hold as the number of loci and the number of sampled individuals become large. The use of the standard χ^2 approximation (*i.e.*, comparing two times the log-likelihood ratio to a χ_1^2 distribution) is appropriate for large samples. However, for small samples, especially when the number of loci is small, the χ^2 approximation may not necessarily provide a good approximation to the distribution of the likelihood-ratio test statistic. We therefore performed simulations to investigate the applicability of the large sample approximations for moderate sample sizes. Data sets were simulated assuming samples sizes of $n_1 = 226$, $n_2 = 122$, $c = 146$, and $N = 5100$. This corresponds to the sample size in the observed data, which are analyzed in the subsequent section. The number of loci ($n = 6$) and the allele frequencies were also chosen to match the values observed in the humpback whale data. The results of the simulations can be found in Figure 5. Note the very close fit between the simulated distribution of likelihood-ratio statistics and the χ^2 distribution. It appears that the χ^2 approximation works well even for these limited sample sizes.

An additional set of simulations was made assuming

20% of the sampled potential fathers are paternal sibs to individuals in the offspring generation. Again it appears that the χ^2 distribution provides a close approximation to the distribution of the likelihood-ratio test statistic, especially in the tail of the distribution. At the 5% significance level, the χ^2 approximation provides a critical value of 3.84 and the true value is $\sim$3.98.

APPLICATION TO THE NORTH ATLANTIC HUMPBACK WHALE

North Atlantic humpback whales congregate mainly on shallow breeding grounds in the West Indies during the winter, which constitutes the breeding season (WHITEHEAD and MOORE 1982). Observational and population genetic data strongly suggest that humpback whales observed in the West Indies constitute a single panmictic population (MATTILA *et al.* 1989; CLAPHAM *et al.* 1993; PALSBØLL *et al.* 1997a, 1998). Females give birth to a single calf on average every second year, although longer and shorter birth intervals have been recorded (CLAPHAM and MAYO 1987, 1990; BARLOW and CLAPHAM 1997). The gestation period has been estimated at $\sim$12 months and the calf is weaned toward the end of its first year.

CLAPHAM (1996) described the humpback whale mating system as polygamous, with many attributes of a lek, where males signal by "singing" and compete for access to estrous females. As many as 25 males have been observed to compete for access to a single, presumably estrous, female during the breeding season (MATTILA *et al.* 1989; CLAPHAM *et al.* 1992). Males in these competitive groups can be divided into several roles, as described by CLAPHAM *et al.* (1992): the principal escort, which is the primary escort of the female (termed the nuclear animal); the challenger, the male whale that actively challenges the principal escort for his position; and the secondary escorts, which denote any other whale in the group. Principal escorts and challengers are considered key male roles and are assumed to be more dominant animals than secondary escorts. The secondary escorts are only rarely observed challenging the principal escort. Such competitive groups of males may last many hours and supposedly require a substantial investment by the dominant males (MATTILA *et al.* 1989; CLAPHAM *et al.* 1992) of which the return presumably is a relatively higher proportion of successful paternities.

Our objective here is to estimate and assess the relative difference in reproductive success of the dominant males (principal escorts and challengers, designated group 1) and the subdominant males (the secondary escorts, designated group 2) from genotypic data.

Our analysis focuses on individual humpback whales sampled in the West Indies during the breeding seasons of 1992 and 1993. These samples constitute a subset of the 3060 tissue samples collected either as skin biopsies (PALSBØLL *et al.* 1991) or sloughed skin (CLAPHAM *et al.*
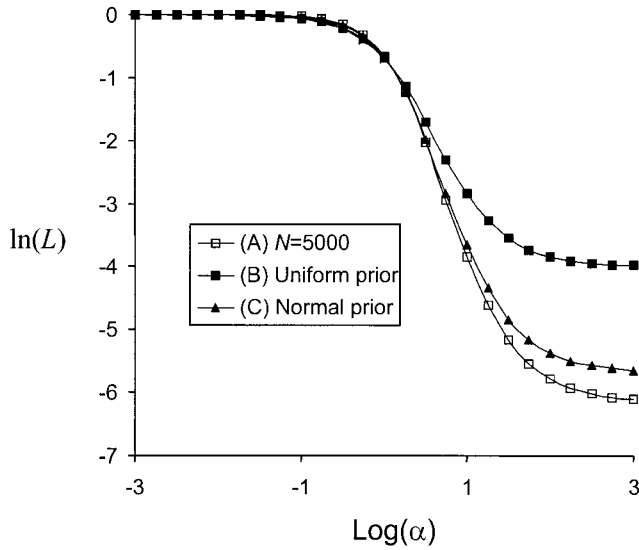
FIGURE 6.—The likelihood surface for α calculated for the Baleen data described in the text. In A it is assumed that $N = 5000$. In B a uniform prior is assumed for N and the integrated likelihood function for α is plotted. In C a normal prior is assumed for N with mean 5000 and standard deviation 1000.

1993) from humpback whales across the North Atlantic Ocean between 1988 and 1995. The genotype at six microsatellite loci and sex were determined for each sample (see PALSBØLL *et al.* 1997a,b and SMITH *et al.* 1999 for details). The microsatellite analyses yielded 2368 unique genotypes among the 3060 samples, each of which was inferred to represent a single individual. The average number of alleles per locus was estimated at 14.5 (PALSBØLL *et al.* 1999). A total of 146 complete mother-calf pairs, as well as 226 males from group 1 and 122 males from group 2, were sampled in 1992 or 1993 on the breeding range among the sample of 2368 unique genotypes. The remaining samples were either collected in different years, on the feeding grounds, or from behavioral classes not relevant to this study, such as pairs and single individuals.

The likelihood function of α is shown in Figure 6A, assuming a population size of 5000 ($= N$) males, which is the most current direct estimate and based on data from 1992 and 1993, the years in which the samples for this study were collected (SMITH *et al.* 1999). The maximum-likelihood value of α is a strictly decreasing function ($\hat{\alpha} = 0$). This result suggests that group 1 (principal escorts and challengers) may have a larger reproductive success than group 2 (the secondary escorts). However, the difference is not statistically significant. A 5% confidence region for α is given by {α : $0 \leq \alpha < 3.1$}. This large confidence interval is a consequence of the flat likelihood surface. The amount of information in the data regarding α is very limited because the number of sampled males and mother-calf pairs is small relative to the overall population size.

To illustrate this problem, we can estimate the ex-

pected number of offspring in the sample from each group (O_s, $s = 1, 2$), conditional on the data, assuming $\alpha = 1$:

$$E(O_s) = \sum_i \frac{\Pr[\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}^{(s)}, I_s(i)]\Pr[I_s(i)]}{\Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}^{(1)}, \mathbf{F}^{(2)})}. \qquad (6)$$

On the basis of the data discussed above, the expected numbers of offspring from males observed from the two groups are 6.26 and 1.93, respectively. In conclusion, the number of expected matches contained in the current sample appears to be too small to provide narrow confidence intervals for α.

ESTIMATION OF EFFECTIVE POPULATION SIZE

In the derivation of the method for paternity assignment described above, it is evident that the likelihood is a function of the number of breeding males. Hence, it is possible to estimate the number of breeding males from the genotypic data. Note that such an estimate of population size is much different from traditional estimates of population sizes based on inbreeding coefficients or similar measures (*e.g.*, KUHNER *et al.* 1995). First, the method provides an estimate of the actual number of potentially breeding males. Population genetic estimates, in contrast, are usually scaled with the mutation rate, which is often an unknown quantity. Second, population genetic estimates are evolutionary estimates, which reflect past events, such as fluctuations in effective population size. The estimate based upon parent offspring genotypes is an estimate of the current male population size, *i.e.*, at the time of sampling.

Assuming independence among offspring, the likelihood function for N can be calculated as

$$L(N) \propto \prod_i \Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{A}, \mathbf{F}, N)$$

$$= \prod_i \left(\frac{(N - n)}{N} \Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{A}) + \frac{1}{N} \sum_{j=1}^n \Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}_j) \right),$$

$$N \geq n. \qquad (7)$$

The likelihood function for N for North Atlantic humpback whales, based on the previously discussed data, is shown in Figure 7. The maximum-likelihood estimate of N is 6540 breeding males ($l = -2128.3571$) and an $\sim$95% confidence interval is given by {N: $3700 \leq N < 17,000$} using parametric bootstrapping. The confidence interval provided by large sample theory is almost identical {N: $3800 \leq N < 16,760$}.

A more direct estimate of the number of male humpback whales on the North Atlantic breeding ground has been obtained by mark-recapture methods using genetic tagging (PALSBØLL *et al.* 1997a). This study yielded a point estimate of males $N_{\text{males}} = 4894$ and a 95% confidence interval of {$3374 \leq N_{\text{males}} < 7123$} (PALSBØLL *et al.* 1997a). The two estimates are quite compatible, but the confidence interval provided by the
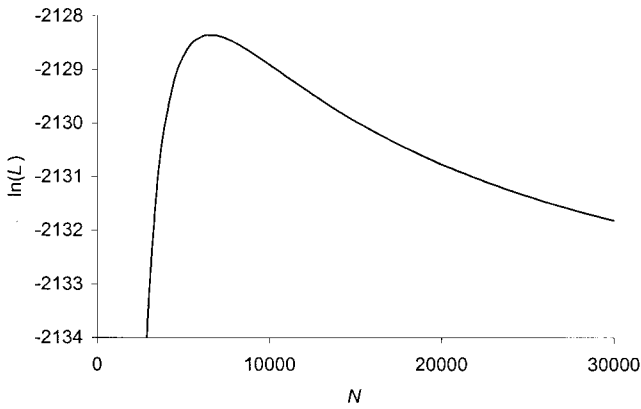
Figure 7.—The likelihood surface for *N*. The likelihood surface is calculated for the Baleen data described in the text.

mark-recapture method is, not surprisingly, considerably narrower than the confidence interval based on the parent-offspring data. Since the assumptions underlying the two estimates are quite different, it is somewhat comforting that the estimates are so similar.

One caveat is that the method assumes that the prior probability of paternity equals $1/N$. If males with a relatively high reproductive success are preferentially sampled, our method will tend to underestimate the male population size. For example, in this study, sampling within the competitive groups was directed toward the dominant males at the expense of the subdominant males. This inherent feature of the sampling design might bias our estimate of *N* toward smaller values.

PATERNITY INFERENCE WHEN THE POPULATION SIZE IS UNKNOWN

In the paternity analysis discussed above, it was assumed that the population size was known. This was a reasonable assumption because of the availability of good census estimates based on mark-recapture methods for the North Atlantic humpback whale population (Smith *et al.* 1999). Unfortunately, the male breeding population size *N* may not be known with great confidence in many cases. In such cases, simulation approaches may be useful when making binary decisions regarding paternity. However, probabilities of paternity may still be desirable, for example, for examining hypotheses regarding the reproductive success of different biologically defined groups. In the following we discuss some methods for calculating these probabilities.

When some (limited) information is available regarding the population size, it may be desirable to take the uncertainty regarding this parameter into account by assuming a prior distribution of the male population size, $[f(N)]$. For example, if a point estimate of *N* with large confidence intervals is available, *N* can be appropriately modeled, for example, as having a normal or

a lognormal distribution, and the posterior probability of paternity can be calculated as

$$\Pr(I_k(i) \mid \mathbf{M}_i, \mathbf{F}, \mathbf{A}) = \int_n^\infty \Pr(I_k(i) \mid \mathbf{M}_i, \mathbf{F}, \mathbf{A}, N) f(N) \, dN. \tag{8}$$

This one-dimensional integral can be evaluated quite easily by standard numerical integration algorithms (*e.g.*, Press *et al.* 1988, pp. 129ff). The density $f(N)$ approximates the true discrete distribution of *N*. Since the integral in Equation 8 and the subsequent equations are evaluated by numerical integration on a grid, there is no practical difference between assuming a discrete and a continuous distribution.

The distribution can be updated using the data of parent and offspring genotypes

$$\Pr(I_k(j) \mid \mathbf{M}_j, \mathbf{F}, \mathbf{A})$$

$$= \frac{\int_n^\infty \Pr(I_k(j) \mid \mathbf{M}_j, \mathbf{F}, \mathbf{A}, N) \prod_{i:i \neq j} \Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{A}, \mathbf{F}, N) f(N) \, dN}{\int_n^\infty \prod_{i:i \neq j} \Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{A}, \mathbf{F}, N) f(N) \, dN}. \tag{9}$$

In this way, the probability of paternity can be calculated using the information regarding population size available in the genetic data from the entire sample. This approach can also be used even if no prior information is available regarding population size. In such cases, it may be reasonable to use a uniform prior for *N*, *i.e.*, to assign equal weight to all possible values of *N*. For most data, it may be necessary to specify a maximum male population size to ensure that the resulting posterior distribution is proper, *i.e.*, $f(N) = 1/(N_{max} - n)$, $n \leq N < N_{max}$.

As a practical approach, it may be computationally simpler to use

$$\Pr(I_k(j) \mid \mathbf{M}_j, \mathbf{F}, \mathbf{A}) \approx \int_n^\infty \Pr(I_k(j) \mid \mathbf{M}_j, \mathbf{F}, \mathbf{A}, N) w_n \, dN,$$

$$w_n = \frac{\prod_i \Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{A}, \mathbf{F}_j, N) f(N)}{\int_n^\infty \prod_i \Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{A}, \mathbf{F}_j, N) f(N) \, dN}. \tag{10}$$

For large samples, Equation 10 should provide a very good approximation. Similarly, inference regarding reproductive success can be performed using the integrated likelihood for α:

$$L_m(\alpha) = \int_n^\infty \prod_{i=1}^k \Pr(\mathbf{O}_i \mid \mathbf{M}_i, \mathbf{F}^{(1)}, \mathbf{F}^{(2)}, N, \alpha) f(N) \, dN. \tag{11}$$

In this way, it is possible to examine hypotheses regarding reproductive success, while incorporating the relevant information from the genetic data regarding the male population size. An example is shown in Figure 6. A uniform prior $f(N) = 1/(20,000 - n)$, $n \leq N < 20,000$ (Figure 6B) or a normal with $\mu = 5000$ and $\sigma = 1000$ (Figure 6C) was used and the likelihood surface was evaluated by numerically integrating Equation 11 on a grid containing 200 grid points. As expected, the
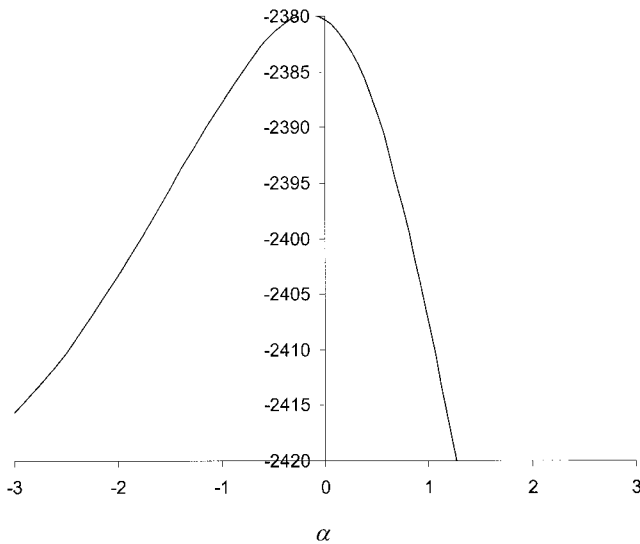
FIGURE 8.—The likelihood surface for α using the fractional paternity method. The likelihood surface is calculated for the Baleen data described in the text.

likelihood surface is more flat when uncertainty regarding N is incorporated into the method because added uncertainty regarding N leads to a loss of statistical power. However, the major features of the likelihood function are retained and the maximum-likelihood estimate of α is zero in all cases.

To illustrate why it is not recommendable to use the fractional-likelihood method when there are unobserved males, we also calculated the likelihood function for α using this method (Figure 8). Note that an estimate of α close to 1 is obtained. Also note that the likelihood function is very peaked, implying that we would have had very strong (false) confidence in this conclusion. Quite intuitively, many males of both group 1 and group 2 would be falsely assigned as parents. Consequently, it would appear as if both groups have similar reproductive success.

DISCUSSION

Estimation of reproductive success in male North Atlantic humpback whales: Using the methods developed in this article, we attempted to test the hypothesis of differential male reproductive success as well as the number of breeding males among North Atlantic humpback whales.

We estimated the relative reproductive success of presumed dominant males and subdominant males sampled on the breeding range in 1992 and 1993. While our sample contained only a small fraction of the total population and thus yielded estimates with wide confidence intervals, our results are in accordance with the hypothesis that dominant males indeed have a relatively higher reproductive success than subdominant whales. The average group size of male competitive groups in

the West Indies during 1992 and 1993 was 4.65 ($n = 289$ groups, 95% confidence interval of ± 0.23; J. Robbins, unpublished results). This implies that the population frequency of subdominant males is only ~30% more than that of the dominant males and thus the dominant males are likely to sire approximately three times more of the calves than the subdominant males. This conclusion is highly tentative, though, as our sample sizes were too small to yield any significant difference in reproductive success between the two groups of males, despite the apparently large difference in the estimate of relative reproductive success.

The average number of alleles per locus (estimated at 14.5, see above) was within the range sufficient for successful parentage assignment, as suggested by our simulation experiments. However, the allele frequencies were far from equal, with an average of 30 and 20% of the alleles at frequencies <0.01 or >0.1, respectively. Maybe more important is the overall proportion of the population that was sampled in this study. The most current abundance estimate for humpback whales in the North Atlantic is 10,600 (SMITH *et al.* 1999). Even though the overall sample of analyzed North Atlantic humpback whales is relatively large (2368), it comprises only 22% of the overall population, of which only a fraction of this sample was for the estimations presented in this study. Given the rather low proportion sampled from the population, the expected number of calves contained in our sample is low, explaining the lack of statistical power in the analysis.

It would be possible to improve the power without increasing the number of sampled individuals. If sufficiently many loci have been sampled, it may be possible to estimate pedigrees and thereby identify all parent-offspring relations among all individuals in the total sample (in this example the 2368 individual humpback whales sampled). Such an approach would greatly increase the number of available parent-offspring pairs without increasing the sample size and may therefore present a practical approach for elucidating the important biological problems investigated in this study.

Abundance estimation of reproductive males from parent-offspring genotypes: The method presented in this study was also used to obtain an abundance estimate of reproductive males. The maximum-likelihood value for the number of breeding males on the North Atlantic breeding range was estimated at 6540, with a 95% confidence interval of 3800–16,760. Our estimate was comparable to the estimate obtained by mark-recapture methods based upon genetic tagging of males (PALSBØLL *et al.* 1997a), which yielded a point estimate of 4890 males and a 95% confidence interval of 3370–7120. While it is not surprising that the confidence interval is much narrower for the latter estimate, it is reassuring that the two estimates are in overall agreement. Interestingly, the lower bound of the 95% confidence interval of 3800 breeding males obtained in this study indicates

a relatively large effective population size of breeding males, which further corroborates the notion that it is unlikely that a few dominant males sire the majority of calves.

Our findings are consistent with the known characteristics of the mating system of this species, in which mature females have a widespread (*i.e.*, nonclustered) distribution (CLAPHAM 1996). Consequently it is difficult for a few dominant males to monopolize and inseminate large numbers of females as observed in other marine mammals (*e.g.*, elephant seals), which leads to a low variance in reproductive success among male humpback whales.

LITERATURE CITED

AMOS, B., J. BARRETT and G. A. DOVER, 1991 Breeding behavior of pilot whales revealed by DNA fingerprinting. Heredity **67:** 49–55.

AMOS, B., C. SCHLÖTTERER and D. TAUTZ, 1993 Social structure of pilot whales revealed by analytical DNA profiling. Science **260:** 670–672.

BARLOW, J., and P. J. CLAPHAM, 1997 A new birth-interval approach to estimating demographic parameters of humpback whales. Ecology **78:** 535–546.

CLAPHAM, P. J., 1996 The social and reproductive biology of Humpback whales: an ecological perspective. Mamm. Rev. **26:** 27–49.

CLAPHAM, P. J., and C. A. MAYO, 1987 Reproduction and recruitment of individually identified humpback whales, *Megaptera novaeangliae*, observed in Massachusetts Bay, 1979–1985. Can. J. Zool. **65:** 2853–2863.

CLAPHAM, P. J., and C. A. MAYO, 1990 Reproduction of humpback whales, *Megaptera novaeangliae*, observed in the Gulf of Maine. Rep. Int. Whaling Comm. Spec. Issue **12:** 171–175.

CLAPHAM, P. J., and P. J. PALSBØLL, 1997 Molecular analysis of paternity shows promiscuous mating in female humpback whales (*Megaptera novaeangliae*, Borowski). Proc. R. Soc. Lond. Ser. B **264:** 95–98.

CLAPHAM, P. J., P. J. PALSBØLL, D. K. MATTILA and O. VASQUEZ, 1992 Composition and dynamics of humpback whale competitive groups determined by molecular analysis. Behaviour **122:** 182–194.

CLAPHAM, P. J., D. K. MATTILA and P. J. PALSBØLL, 1993 High-latitude-area composition of humpback whale groups in Samana Bay: further evidence for panmixis in the North Atlantic population. Can. J. Zool. **71:** 1065–1066.

DEVLIN, B., K. ROEDER and N. C. ELLSTRAND, 1988 Fractional paternity assignment—theoretical development and comparison to other methods. Theor. Appl. Genet. **76:** 369–380.

FOLTZ, D. W., and D. W. HOGLAND, 1981 Analysis of the mating system in the black-tailed prairie dog (*Cynomys ludovicianus*) by likelihood of paternity. J. Mamm. **62:** 706–712.

KUHNER, M. K., J. YAMATO and J. FELSENSTEIN, 1995 Estimating effective population size and mutation rate from sequence data using Metropolis-Hastings sampling. Genetics **140:** 1421–1430.

MARSHALL, T. C., J. SLATE, L. E. B. KRUUK and J. M. PEMBERTON, 1998 Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. **7:** 639–655.

MATTILA, D. K., P. J. CLAPHAM, S. K. KATONA and G. S. STONE, 1989 Population composition of humpback whales, *Megaptera novaeangliae*, on Silver Bank, 1984. Can. J. Zool. **67:** 281–285.

MEAGHER, T. R., 1986 Analysis of paternity within a natural population of *Chamaelirium luteum*. I. Identification of most-likely male parents. Am. Nat. **128:** 199–215.

MEAGHER, T. R., and E. A. THOMPSON, 1986 The relationship between single parent and parent pair genetic likelihoods in genealogy reconstruction. Theor. Popul. Biol. **29:** 87–106.

MEAGHER, T. R., and E. A. THOMPSON, 1987 Analysis of parentage for naturally established seedlings within a population of *Chamaelirium luteum* (Liliaceae). Ecology **68:** 803–812.

PALSBØLL, P. J., F. LARSEN and E. SIGURD HANSEN, 1991 Sampling of skin biopsies from free-ranging large cetaceans in West Greenland: development of new biopsy tips and bolt designs. Rep. Int. Whaling Comm. Spec. Issue **13:** 71–79.

PALSBØLL, P. J., J. ALLEN, M. BÉRUBÉ, P. J. CLAPHAM, T. P. FEDDERSEN *et al.*, 1997a Genetic tagging of humpback whales. Nature **388:** 676–679.

PALSBØLL, P. J., M. BÉRUBÉ, A. H. LARSEN and H. JØRGENSEN, 1997b Primers for the amplification of tri- and tetramer microsatellite loci in cetaceans. Mol. Ecol. **6:** 893–895.

PALSBØLL, P. J., P. J. CLAPHAM, H. JØRGENSEN, F. LARSEN, D. K. MATTILA *et al.*, 1998 The value of parallel analysis of uni- and bi-parental inherited loci: the North Atlantic humpback whale (*Megaptera novaeangliae*), pp. 426–430 in *Molecular Tools for Screening Biodiversity: Plants and Animals*, edited by A. KARP, P. G. ISAAC and D. S. INGRAM. Chapman & Hall, London.

PALSBØLL, P. J., M. BÉRUBÉ and H. JØRGENSEN, 1999 Multiple levels of single-strand slippage at cetacean tri- and tetranucleotide repeat microsatellite loci. Genetics **151:** 285–296.

PRESS, W. H., S. A. TEUKOLSKY, W. T. VETTERLING and B. P. FLANNERY, 1988 *Numerical Recipes in C*. Cambridge University Press, Cambridge, UK.

ROEDER, K., B. DEVLIN and B. G. LINDSAY, 1989 Application of maximum likelihood methods to population genetic data for the estimation of individual fertilities. Biometrics **45:** 363–379.

SMITH, T. D., J. ALLEN, P. J. CLAPHAM, N. FRIDAY, P. S. HAMMOND *et al.*, 1999 An ocean-basin-wide mark-recapture study of the North Atlantic humpback whale (*Megaptera novaeangliae*). Mar. Mamm. Sci. **15:** 1–32.

SMOUSE, P. E., and T. R. MEAGHER, 1994 Genetic analysis of male reproductive contributions in *Chamaelirium luteum* (L.) Gray (Liliaceae). Genetics **136:** 313–322.

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

THOMPSON, E. A., 1976 Inference of genealogical structure. Soc. Sci. Inf. **15:** 173–188.

WHITEHEAD, H., and M. J. MOORE, 1982 Distribution and movements of West Indian humpback whales in winter. Can. J. Zool. **60:** 2203–2211.

Communicating editor: M. W. FELDMAN