

COMMENT

High levels of statistical uncertainty in 'gametic' recapture estimates of male abundance in humpback whales

Per J. Palsbøll^{1,*}, Martine Bérubé¹, Eric C. Anderson², Kevin K. Dunham²

¹Department of Environmental Science, Policy and Management, University of California Berkeley, 137 Mulford Hall, Berkeley, California 94720, USA

²NOAA Southwest Fisheries Science Center, Santa Cruz Laboratory, 110 Shaffer Road, Santa Cruz, California 95060, USA

Garrigue et al. (2004) compared organismal and gametic 'mark-recapture' estimates of male abundance of humpback whales *Megaptera novaeangliae* to assess whether the population wintering off New Caledonia constituted a self-contained, reproductively autonomous population, i.e. whether or not its members were inter-breeding with animals from other areas. The approach taken by Garrigue et al. (2004) attempts to address an old and common, but difficult, issue in ecology and management. Usually the delineation of independent breeding populations is inferred from the spatial distribution of re-sightings of marked animals (e.g. Baker et al. 1986), or in the case of genetic data, by homogeneity tests of allele frequencies among areas (Moritz 1994). When only a single area is sampled (as in Garrigue et al. 2004), then genotypic data can be employed to test whether genotype frequencies deviate significantly from the expected Hardy-Weinberg proportions expected in a single panmictic population (Hardy 1908). However, such an assessment has low statistical power. Garrigue et al. (2004) found that the organismic estimate of male abundance ($N = 382$) was sufficiently close to the gametic estimate of male abundance ($N = 379$) to argue that the humpback whales wintering off New Caledonia constitute a single autonomous, panmictic breeding population. While other previous studies have estimated male abundance in a similar manner (i.e. from genetic paternity estimation; Nielsen et al. 2001, Pearse et al. 2001, Skaug 2001, Ramakrishnan et al. 2004), the purpose of the estimation undertaken by Garrigue et al. (2004) is novel and may constitute a more powerful manner by which to assess whether a

single sample constitutes an autonomous breeding population. However, there is a potentially serious issue with this approach that was not addressed by Garrigue et al. (2004), namely the possibility that the observed result could have occurred even if some or many 'outside' males (from another breeding population) were mating with New Caledonia females.

We conducted computer demographic simulations to assess how large the proportion of outbreeding females could have been while still yielding the result observed by Garrigue et al. (2004). Our simulations were based upon the demographic parameters reported by Rosenbaum et al. (2002) using an unpublished multi-population version of the program *spip* (Anderson & Dunham in press). The maximum age of an individual was set at 60 yr. The probability of survival each year was set at 0.626 for 1 yr old calves, 0.825 for 2 to 5 yr old animals, and 0.96 for animals between 6 and 60 yr old. Both sexes were assumed to reach sexual maturity at an age of 6 yr and the probability of an individual reproducing each year was set at 0.8. A binary offspring distribution was used with a reproduction inhibition of 2 yr. The simulations thus matched 1 yr of pregnancy followed by 1 yr of lactation and weaning.

We simulated 2 populations simultaneously to represent the populations given in Garrigue et al. (2004): a smaller population averaging 660 sexually mature individuals (330 males) to represent the New Caledonia population, and a larger population averaging 2500 mature individuals (1250 males) to represent the larger Australian population, which New Caledonia female humpback whales could potentially outbreed with.

*Email: palsboll@berkeley.edu

The proportion of males, U , from the large Australian population available to New Caledonia females was varied from 0 to 0.5 in steps of 0.1. For each value of U , 500 simulations were conducted. Initial numbers of different age groups were set to those expected under demographic equilibrium, and the populations were simulated forward in time. At $T = 100$ yr during each simulation all 1 yr old calves and approximately 35% of sexually mature males from the representative New Caledonia population were genotyped. We simulated the genotyping of 35% of the males, because that is the estimated proportion (133/382) of New Caledonia males sampled by Garrigue et al. (2004). No data were collected from the larger Australian population. From the genotyped 1 yr old calves, 16 were randomly chosen and compared with the set of genotyped males to see whether any of the males could be identified as fathers. The number of father–calf identifications was recorded and a new simulation was started. The father–calf identifications were made without error (as if an infinite amount of genetic data were available).

The proportion of times out of 500 simulations that 5 or more calves were assigned to males in the New Caledonia population given different degrees of availability of males from other parts of the range (i.e. the larger Australian population) was; $p_{\text{zero}} = 0.59$; $p_{0.1} = 0.31$; $p_{0.2} = 0.13$; $p_{0.3} = 0.082$; $p_{0.4} = 0.038$; and $p_{0.5} = 0.016$. The subscript denotes the fraction of males from the larger Australian population that are available for New Caledonia females, e.g. 0.2 implies that 250 (1250×0.2) Australian males were available to New Caledonia females. The simulations show a relatively high probability ($p = 0.082$) of the observed result even if there is a large contribution (i.e. around $375 = 1250 \times 0.3$) of breeding males from outside New Caledonia. In other words, even if over 50% [$375/(375 + 330)$] of the males available to New Caledonia females come from outside the New Caledonian population, the result observed by Garrigue et al. (2004) is expected to occur >5% of the time.

This result is not surprising, and it is straightforward to verify it analytically, as follows: (1) We adopt the notation of Garrigue et al. (2004), using n_1 to denote the number of non-calf males sampled and genotyped, n_2 for the number of calves sampled and genotyped, m for the number of inferred paternities among the n_2 calves, and N_m for the total number of reproductively mature males from the New Caledonia population, and we let A denote the number of males from the Australian population (or any other non-New Caledonia population) of humpback whales that have mating access to females from the New Caledonia population. (2) We note that randomly sampling the n_2 calves and identifying their fathers can be modeled as a simple process of randomly sampling n_2 fathers from the pop-

ulation of reproductively mature males accessible to New Caledonia females. The probability, q , that each father thus sampled is one of the n_1 previously genotyped males is just the proportion of previously genotyped males amongst all mature males accessible to the females. Hence, $q = n_1/(N_m + A)$, and m has a binomial distribution of n_2 trials, each with a success probability of q .

Given the above setup, the probability of observing $m = 5$ or more calves with inferred paternity from the sampled males is just the probability, given q , that $m \geq 5$. In Fig. 1A we plot this probability for different values of A using the values of $n_1 = 133$, $n_2 = 16$, and $m = 5$, as in Garrigue et al. (2004). The lower x-axis shows the number of Australian males (or males from any other 'outside' population) available as mates to the New Caledonia females. The solid line gives the relationship between the probability of observing 5 or more calves with paternity belonging to one of the 133 genotyped males (on the y-axis) and the number of Australian males, when N_m is equal to the value of 382, as estimated by Garrigue et al. (2004). The upper and lower dotted lines show the same for the cases of $N_m = 554$ and $N_m = 220$, the upper and lower 95% confidence interval endpoints of Garrigue et al.'s estimate of N_m , respectively. The upper x-axis gives the proportion of all males (New Caledonian or otherwise) available to New Caledonia females that are not from the New Caledonia population. Hence, we see that even if 34% of the males available to the New Caledonia females are from the Australian population, there is still a 30% chance of observing 5 or more calves with fathers from the group of genotyped males. The open circles on the plot show where the simulation results fall when plotted on the upper x-axis against the y-axis. These points fall below the solid curve, because our simulated male population size (averaging 330) was smaller than the 382 estimated by Garrigue et al. (2004). However, the conclusion remains the same: there is insufficient evidence from the gametic mark-recapture method alone, to conclude that the New Caledonia population is reproductively autonomous.

It is worthwhile asking how much data would be necessary for a gametic recapture inference to provide conclusive evidence that a population is reproductively autonomous. It is apparent that the most valuable increase in sampling effort would be to increase the proportion of genotyped males (i.e. to increase n_1/N_m). If n_1/N_m remains as low as $133/382 = 35\%$, then simply increasing sample sizes does not achieve a strong improvement. This is shown in Fig. 1B which depicts a situation where $n_1 = 1333$, $n_2 = 160$, $m = 50$, and $N_m = 3820$ (as in the New Caledonia whale scenario, but all samples and populations have been increased 10 \times). In this case, n_1/N_m is still 35%, but it is

still possible, with a probability of 10%, to observe $m = 50$ inferred paternities for 160 calves, even when 20% of the males available to reproductive females are from an 'outside' population.

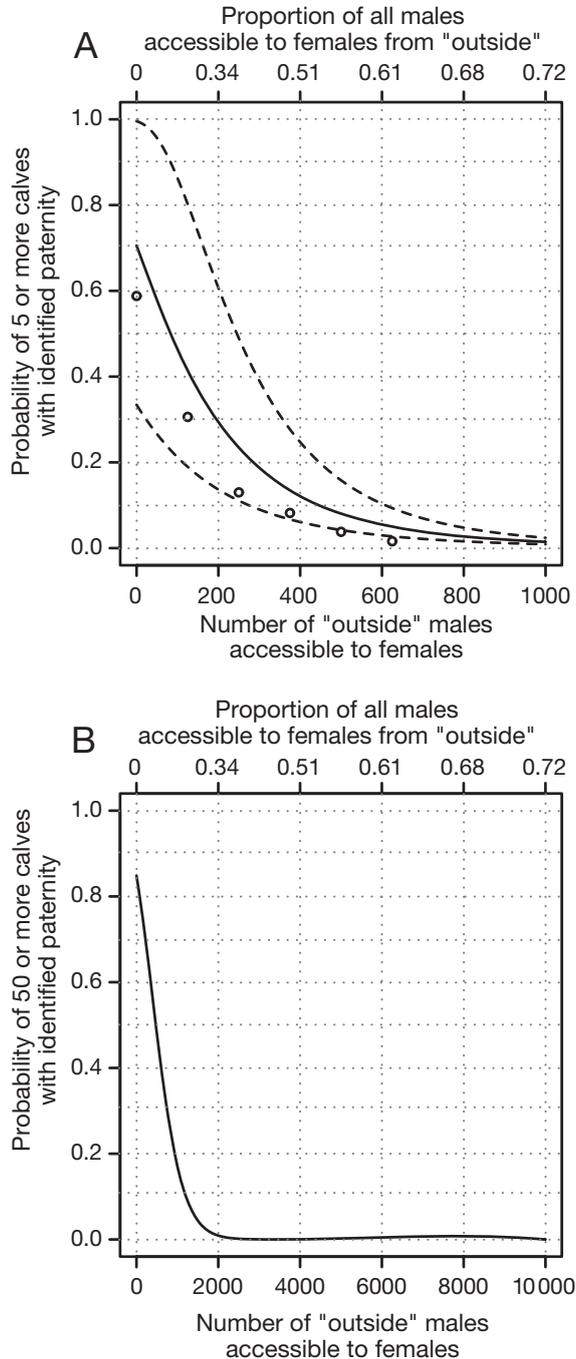


Fig. 1. *Megaptera novaeangliae*. Probability of observing m or more paternities in a sample of n_2 calves as a function of the number of males from outside the population that are available to females. (A) Sample sizes and observed m are set to those in Garrigue et al. (2004). (B) Hypothetical scenario involving population and sample sizes 10 \times greater than in (A)

In conclusion, our simulations and analyses suggest that it may be premature to conclude that the New Caledonia winter ground of humpback whales harbors an autonomous population. While the approach suggested by Garrigue et al. (2004) may help to delineate autonomous populations, it appears as if substantially more data are required to enable a statistically rigorous conclusion.

There are a few additional aspects, which may affect Garrigue et al.'s (2004) and our conclusions. For instance, as the number of 'outside' males increases, so should the organismic mark-recapture estimate of abundance. In other words, both estimates of abundance will increase in parallel. Finally, both Garrigue et al. (2004) and we assumed that all calves could be assigned with 100% certainty to a sampled father. This is highly unlikely when only 9 loci are analyzed (see Palsbøll 1999 for simulations of statistical power). The inclusion of incorrect paternity assignments will result in an underestimate of male abundance.

LITERATURE CITED

- Anderson EC, Dunham KK (in press) spip 1.0: a program for simulating pedigrees and genetic data in age-structured populations. *Mol Ecol Notes*
- Baker CS, Herman LM, Perry A, Lawton WS and 7 others (1986) Migratory movement and population structure of humpback whales (*Megaptera novaeangliae*) in the central and eastern North Pacific. *Mar Ecol Prog Ser* 31: 105–119
- Garrigue C, Dodemont W, Steel D, Baker CS (2004) Organismal and 'gametic' capture-recapture using microsatellite genotyping confirm low abundance and reproductive autonomy of humpback whales on the wintering grounds of New Caledonia. *Mar Ecol Prog Ser* 274: 251–262
- Hardy GH (1908) Mendelian proportions in a mixed population. *Science* 28:49–50
- Moritz C (1994) Defining 'Evolutionary Significant Units' for conservation. *Trends Ecol Evol* 9:373–375
- Nielsen R, Mattila DK, Clapham PJ, Palsbøll PJ (2001) Statistical approaches to paternity analysis in natural populations and applications to the North Atlantic humpback whale. *Genetics* 157:1673–1682
- Palsbøll PJ (1999) Genetic tagging: contemporary molecular ecology. *Biol J Linnean Soc* 68:3–22
- Pearse DE, Eckerman CM, Janzen FJ, Avise JC (2001) A genetic analogue of 'mark-recapture' methods for estimating population size: an approach based on molecular parentage assessments. *Mol Ecol* 10:2711–2718
- Ramakrishnan U, Storz JF, Taylor BL, Lande R (2004) Estimation of genetically effective breeding numbers using a rejection algorithm approach. *Mol Ecol* 13:3283–3292
- Rosenbaum HC, Weinrich MT, Stoleson SA, Gibbs JP, Baker CS, DeSalle R (2002) The effect of differential reproductive success on population genetic structure: correlations of life history with matriline in humpback whales of the Gulf of Maine. *J Heredity* 93:389–399
- Skaug HJ (2001) Allele-sharing methods for estimation of population size. *Biometrics* 57:750–756