Behav Ecol Sociobiol (1987) 21: 35–45

Behavioral Ecology and Sociobiology

The use of genetic markers to estimate the frequency of successful alternative reproductive tactics

David F. Westneat, Peter C. Frederick*, and R. Haven Wiley
Department of Biology, University of North Carolina, Chapel Hill, NC 27514, USA

Received September 10, 1986 / Accepted February 20, 1987

Summary. This paper outlines a method for estimating rates of successful alternative reproductive tactics from parental exclusions known through the use of genetic markers. We review a method for calculating the probability of excluding a putative father when he is not the actual father. We adapt this method to model two mating tactics of concern to sociobiologists: extrapair copulations (EPCs) and intra-specific egg parasitism (egg-dumping). Four different types of parental exclusions are possible (both male and female, male only, female only, and ambiguous). The two models predict different proportions of each type of exclusion. Models are also generated for the case when the putative mother's or father's genotypes are not available.

We used parental exclusions from an electrophoretic study of indigo buntings (Westneat 1987b) to demonstrate these methods. The distribution of parental exclusions in the buntings departed significantly from the predictions of the egg-dumping model, but agreed well with those of the EPC model. The probability of detection for the EPC model (0.401) was then used to estimate the actual rate of extra-pair fertilizations (0.421 of all the young sampled). We present a method for calculating a confidence interval on this estimate, which ranged from 0.247 to 0.659. We concluded that these methods will allow the quantitative study of the success of alternative reproductive tactics in a wide variety of species.

Introduction

Individuals of many species pursue a variety of reproductive tactics (Cox and Le Boeuf 1977; Oring 1982; McKinney et al. 1984; Davies 1985; Dunbar 1985). For instance, among birds, both intraspecific egg parasitism or egg-dumping (laying eggs in another individual's nest) and extra-pair copulations (EPCs, copulations in which at least one of the participants is paired to a third individual) have been documented in a wide variety of species (Yom-Tov 1980; Ford 1983; Brown 1984; Gowaty and Karlin 1984; McKinney et al. 1984; Gavin and Bollinger 1985; Frederick 1987; Westneat 1987a, b). In order to estimate the frequency of extra-pair fertilizations or successful egg parasitism field researchers have often relied solely on observational methods (Gladstone 1979; Birkhead et al. 1985; Frederick 1985).

For a number of reasons, behavioral observations usually cannot provide accurate estimates of the number of young resulting from these reproductive tactics. First, even intensive observations under the best conditions cannot produce a complete record of behavior (e.g. Werschkul 1982; Frederick 1987; Westneat 1987a). Second, reproductive tactics are likely to differ considerably in conspicuousness, so that estimates of reproductive success based on behavioral observations are likely to be biased (Fitch and Shugart 1984). Third, in species lacking an intromittent organ (such as most birds), copulation may not reliably indicate insemination (Fisher 1971). Fourth, the effects of sperm competition, sperm precedence, and optimal timing of fertilization are poorly known for most animals (Cheng et al. 1983; Smith 1984). Without this information, it is not possible to assign the parentage of many offspring to particular individuals. Fifth, female receptivity to extra-pair matings varies considerably among individuals and species, and it is unknown what effect these differences have on fertilization rates (Cox and Le Boeuf 1977; Van Tienhoven 1983; Fitch and Shugart 1984; Frederick 1985). Finally, egg-dumping with removal of the host's egg (as in inter-specific parasites; Wyllie

* *Current address:* Department of Wildlife, 118 Newins-Ziegler Hall, University of Florida, Gainesville, Fl 32611, USA

Offprint requests to: D.F. Westneat, Department of Genetics and Development, Emerson Hall, Cornell University, Ithaca, NY 14853, USA

1981), would impede estimation of the rates of egg-dumping.

Genetic markers, such as plumage polymorphisms (Burns et al. 1980) and electrophoretically distinct allozymes (Sherman 1981; Mock 1983), provide a potentially more accurate method for measuring the frequency of successful alternative (or conditional) reproductive tactics. The latter technique has been used successfully in a variety of organisms: ground squirrels (Foltz and Hoogland 1981; Hanken and Sherman 1981), mice (Rasmussen 1964; Foltz 1981), bats (McKracken and Bradbury 1977), rhesus monkeys (Duvall et al. 1976; Curie-Cohen et al. 1983) and several species of birds (Gowaty and Karlin 1984; Gavin and Bollinger 1985; Joste et al. 1985; Mumme et al. 1985; Westneat 1987b).

Genetic markers can be used to assign the parentage of individual offspring to a particular adult only when there is a large number of independent and polymorphic loci (see Silver 1982). The more common use of these markers is to *exclude* a putative parent as the actual parent of the offspring. The distinction between these uses is paramount (Chakraborty et al. 1974; Barrowclough et al. 1985). Throughout this paper, we will concentrate on the use of genetic markers for exclusion and not for inclusion.

It is also important to realize that genetic markers can only reveal a fraction of what we shall call stray genes (genes in offspring from individuals other than the putative parents), because the genotypes of putative and actual parents will in some cases be the same. Thus estimates of the frequencies of stray genes will be consistently underestimated by genetic markers. If genetic markers are used to estimate the frequencies of young that come from successful alternative reproductive tactics, then one must take into account the probabilities of excluding one or both of the putative parents.

A large body of literature on calculating the probability of exclusion in humans (Neel and Schull 1954; Chakraborty et al. 1974; for reviews see Silver 1978, 1982), although primarily oriented toward specialized legal issues, can be adapted to sociobiological applications. Separately, Wrege and Emlen (1987) have made a similar calculation. In this paper we review the general method for estimating the probability of detection of stray genes. We will adapt this method for an examination of two alternative reproductive tactics that are of concern to many sociobiologists: extra-pair fertilizations and intra-specific egg-dumping. We discuss ways in which cases of parental exclusions can be used to identify which of these behavioral causes of stray genes occurs in a population. We also show how the probability of detection can be used to estimate the frequencies of stray genes, and we present a formula for calculating confidence intervals on these estimates.

In order to demonstrate the use of these methods we apply them to electrophoretic data collected by Westneat (1987b) during a study of mating patterns of indigo buntings (*Passerina cyanea*). These data are intended as examples; a full presentation of the raw data is published elsewhere (Westneat 1987b).

Basic model

The probability of detection depends on the genotypes of the putative parents and the actual parents. The probability of detection (d) for a single genetic locus is the probability of detection (e_i) for each arrangement of putative and actual parental genotypes weighted by the probability of that arrangement (a_i) summed over all arrangements, or

$$d = \sum_{i=1}^{n} e_i a_i \tag{1}$$

The simplest model for our purposes applies to a single locus with two alleles as genetic markers. The model assumes that allelic frequencies are in Hardy-Weinberg equilibrium, dominance of one allele is negligible, and all possible parents are picked randomly from the population. The effects of violations of these assumptions on the calculation will be discussed later.

EPC model

This model calculates the probability of detection when the female has mated with a male other than the putative father. Suppose a male with genotype AA pairs with a female with genotype AA. All of their offspring would also be AA. However, suppose the female actually copulated with another male of genotype AB. This mating would only be detected if the offspring turned out to be AB, which has a probability of 0.5. Note that without the assumption that the female is the mother, we cannot tell which parent is excluded. The probability of detection of 0.5 must be weighted by the probability of the mating arrangement. This probability is based on the Hardy-Weinberg combinations of allele frequencies for each of the three genotypes (i.e. $AA = p^2$, $AB = 2pq$, $BB = q^2$). In this example, the probability of the mating arrangement is $p^2 \times p^2 \times 2pq$, or $2p^5q$, and so the probability of detection (0.5) multiplied by the probability of the arrangement is p^5q. In the simple two allele model there are 27 different possible mating arrangements (3 adults and 3 genotypes). Additional examples of mating arrangements are shown in Table 1. Thus for a single locus with two alleles, the overall probability of detection (d) for each locus is the sum of the weighted probabilities of detection over all matings,

$$p^5q + 3p^4q^2 + 4p^3q^3 + 3p^2q^4 + pq^5 \tag{2}$$

which simplifies to

$$pq(1 - pq), \tag{3}$$

Table 1. Selected examples of the calculation of the probability of detection for both models with a two-allele locus. The probability of detection (e_i) for each mating arrangement (probability of occurrence is a_i) is partitioned into the probabilities for each type of exclusion

EPC model

Putative father	Mother	Actual father	e_i					a_i
			Both	Male only	Female only	Ambiguous	Total	
AA	AA	AA	–	0	–	0	0	p^6
		AB	–	0	–	0.5	0.5	$2p^5q$
AA	BB	AB	–	0.5	–	0	0.5	$2p^3q^3$

Egg-dumping model

Putative father	Mother	Actual father	Mother	e_i					a_i
				Both	Male only	Female only	Ambiguous	Total	
AA	AA	AA	BB	0	0	0	1.0	1.0	p^6q^2
		AB	BB	0.5	0	0	0.5	1.0	$2p^5q^3$
AA	BB	AB	AB	0	0.25	0.25	0	0.5	$4p^4q^4$

where p and q ($p+q=1$) are the frequencies of the alleles (Neel and Schull 1954).

We should note here that our method of tallying arrangements of genotypes differs slightly from the method of others. For example, Neel and Schull (1954) focused on the possible genetic identity of the second parent given a particular parent-child combination. Wrege and Emlen (1987) have tallied the offspring genotypes that exclude particular putative parents without specifying the actual parent(s). We have focused on the probability of exclusion with two putative parents and one (or two) actual parents. Our equation for a two-allele locus reduces to Neel and Schull's expression and is presumably a generalization of Wrege and Emlen's (1987) method.

The probability of detection based on any one locus is highest when the allele frequencies are equal and declines as one allele approaches a frequency of 1.0. For any single two-allele locus, the maximum probability of detecting a case of successful extra-pair copulation occurs when $p=q$, and is 0.185.

This simple model can be expanded to calculate probabilities of detection based on loci with more than two alleles (Wiener et al. 1930; Silver 1978). In general the procedure is the same as for two alleles. Additional alleles increase the number of arrangements of genotypes exponentially and thus greatly increase the complexity of the equation for (d). For example, for three alleles there are 216 (3 adults and 6 possible genotypes) and for four alleles 1000 (3 adults and 10 possible genotypes) arrangements of genotypes. A general equation for multiple alleles was developed by Chakravarti and Li (1983). The probability of detection (d) is given by

$$1-2a_2+a_3+3(a_2a_3-a_5)-2(a_2^2-a_4) \quad (4)$$

where $a_n=\sum_{i=1}^{k} p_i^n$, p_i is the frequency of the ith allele, and k is the number of alleles. In the case of two alleles, this equation reduces to equation (3).

Additional alleles increase the probability of detection the most when allele frequencies for a locus are equal (Fig. 1), and only slightly if their frequencies are strongly skewed towards one allele (Wiener et al. 1930; Silver 1978).

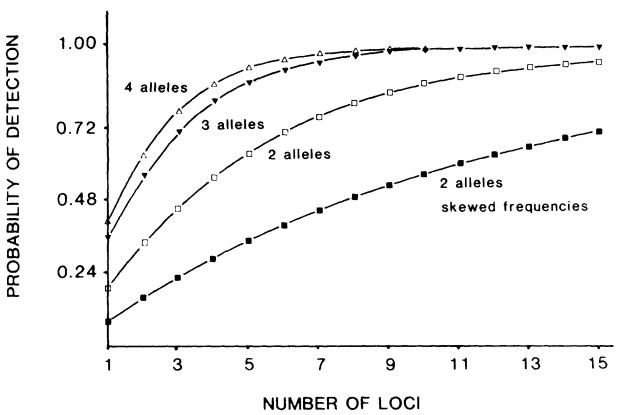

Fig. 1. The effect of the number of loci on the probability of detection (EPC Model) for loci with different numbers of alleles. All alleles are of equal frequency except for the curve labelled 2 alleles (skewed frequency) where $p=0.9$ and $q=0.1$. Equations (4) and (5) were used to generate the curves

If two or more independently segregating loci are used for parental exclusion, the overall probability of detection, D, is one minus the product of one minus the probability of detection at each locus (d_i; Wiener et al. 1930; Chakraborty et al. 1974), or

$$1-\prod_{i=1}^{n}(1-d_i) \quad (5)$$

The more loci, the higher the probability of detection, although each additional locus adds successively less to the overall probability of detection (Fig. 1; Chakraborty et al. 1974). Wrege and Emlen (1987) give a formula for combining the (d_i) for several loci. Their equation reduces to the above.

If one of the putative parents' genotypes is unknown, then the probability of detection is much lower. If the putative father's genotype is unknown, then the probability of detection

Table 2. Equations for the probability of detection of each type of exclusion under the EPC and egg-dumping models for a two allele locus

Type of exclusion				
	Both male and female	Male only	Female only	Ambiguous exclusion
EPC model	0	$2p^2q^2$	0	$pq(p^3+q^3)$
Egg dumping model	$p^2q^2(p^2+q^2)$	$p^2q^2(p^2+4pq+q^2)$	$p^2q^2(p^2+4pq+q^2)$	$2pq(p^4+q^4)$

is quite small. Several authors have generated models for estimating rates of *mixed parentage* in species in which both the putative and actual fathers are unknown (Birdsall and Nash 1973; Foltz 1981; Hanken and Sherman 1981). In these models, the general method is to discover how often more than one male has fertilized a female's litter. Since there is no putative father to exclude, these models calculate the probability of detecting three or more paternal alleles in the litter. These models therefore apply specifically to species in which females are not associated with a particular male and copulate with more than one male.

Using a slight modification of the EPC model, we have derived an equation to describe the probability of detection when the genotype of the mother is unknown. Because the female's genotype is unknown, stray genes will be detected only if the offspring's genotype excludes the male (see Tables 1 and 2). As before, equation (1) is used, but in this case most of the probabilities of detection (e_i) for each mating arrangement are less than in the general EPC model. For two alleles, equation (1) gives a sum that reduces to

$$2p^2q^2 \tag{6}$$

As in the case in which the putative father is unknown, the probability of detection is considerably reduced (by at least 33%) if the genotype of the female is unknown.

Egg-dumping model

The EPC model discussed above assumes that the putative mother is the actual mother of the offspring. However, in many species, the putative mother is not always the actual mother. For example, in some species adoption has been observed (Pierotti 1980; Curie-Cohen et al. 1983) and in birds intraspecific nest parasitism (egg-dumping) also occurs (Yom-Tov 1980; Brown 1984; Gowaty and Karlin 1984; Frederick 1985; Emlen and Wrege 1986). Although maternal exclusion is not directly considered in the literature on parental exclusions in humans, some of the same methods apply.

The EPC model can be adapted to calculate the probability of detection if egg dumping is assumed to be the only source of stray genes. This model makes the same genetic assumptions as the EPC model, but in addition assumes that the parasitic female has mated with some male other than the host female's mate. If the parasitic female mated exclusively with the same male (as in a polygynous relationship; quasi-parasitism; Emlen and Wrege 1987) then the probability of detecting the dumped egg is identical to the probability from the EPC model.

The probability of detecting a dumped egg is calculated in the same manner as in the case of EPCs, by summing the probability of detection (e_i) times the probability of the mating arrangement (a_i) over all of the possible mating arrangements (for examples, see Table 1). For a two allele locus, there are 81 arrangements (4 adults, 3 genotypes). The solution to equation (1) eventually reduces to

$$pq(2p^4+3p^3q+8p^2q^2+3pq^3+2q^4) \tag{7}$$

where p and q are allele frequencies. At any given combination of allele frequencies, this probability is higher than the probability of detecting an extra-pair fertilization because alleles excluding putative parents can come from two individuals instead of one as in the case of EPCs. For a single, two-allele locus the maximum probability of detection is 0.273 (when $p=q$).

Equations for loci with more than two alleles can be generated. We have an equation for three alleles, but it is too cumbersome to present here. To our knowledge a general equation for multiple alleles similar to equation (4) has not been derived.

Comparison of the EPC and egg-dumping model

Although most types of exclusions do not allow identification of the cause of the stray genes in particular cases, extra-pair fertilizations and egg-parasitism do produce different distributions of types of exclusions and thus permit some evaluation of the causes of stray genes in the population. If stray genes come only from EPCs, then we expect that only two types of exclusions will be observed; the genotype of the offspring can either be unambiguously incompatible with the putative father (for instance if the putative father is AA and the offspring is BB) or incompatible with either the putative father or female but not both (ambiguous; if the male is AA, female AA, and the offspring is AB). Exclusion of the male and ambiguous exclusion also occur in the case of egg-dumping. However, egg-dumping should also result in unambiguous exclusion of the putative mother and, if the female was inseminated by a different male, in simultaneous exclusion of both putative parents (putative parents both AA, offspring BB).

The two models predict different distributions of these four types of exclusion. The equations for calculating the probabilities of each type of exclusion under each model are shown in Table 2. These equations were constructed by examining each mating arrangement and partitioning the total probability of detection from that arrangement (e_i) into the probabilities for each type of exclusion (see Table 1). Although we have equations for the types of exclusions for both models in the case of three alleles, we have not presented them because they contain a large number of terms and are quite cumbersome. A Basic computer program containing these additional equations is available from the authors.

To illustrate the uses of these models for investigation of the behavioral causes of stray genes and for estimation of the rates of stray genes, we use data collected by Westneat (1987b) on a marked population of indigo buntings near Niles in south-

western Michigan studied from May through August in 1983 and 1984. Small biopsies of the pectoralis muscle of each adult were removed upon capture and from each offspring at age 5–7 days (for details of the technique see Baker 1981; Westneat et al. 1986). Muscle samples were then analyzed by Westneat (1987b) by standard techniques for starch-gel electrophoresis.

Results

To illustrate the calculation of probabilities of detection and frequencies of stray genes, we analyze data from one year (1984). Data from 1983 are also used in an analysis of the distribution of types of exclusion. Nine polymorphic loci were found in 1984 (Table 3). All but one of the loci fit Hardy-Weinberg distributions of genotype frequencies (total of 17 G-tests in two years on adult gene frequencies only; Sokal and Rohlf 1969). The locus (PGI) that did not fit Hardy-Weinberg in 1984 was very close to Hardy-Weinberg expectations in 1983. Heterozygotes exhibited banding patterns consistent with the known subunit structure for each enzyme (Avise et al. 1980; C.F. Aquadro, personal communication). No evidence of linkage disequilibrium between loci was found in pairwise *G*-tests on the observed distribution of individuals with 0, 1, or 2 heterozygous loci for all 36 combinations of loci in 1984 (for details of these analyses, see Westneat 1987b).

To calculate the probability of detecting an extra-pair fertilization in 1984, we substituted the allele frequencies (Table 3) for each locus into the EPC model with the appropriate number of alleles. The probabilities of detection for each locus were combined using equation (5) to provide an overall probability of detection of 0.401 (Table 3). Separately, the probability of detecting each of the four types of exclusion (male excluded, female excluded, both excluded, ambiguous exclusion) was calculated for each locus. The probabilities within each type were then combined using equation (5) to give probabilities of detecting each type of exclusion over all loci (Table 4). The probability of detecting a dumped egg was also calculated for each type of exclusion at each locus. The rarest alleles in four and five allele loci were lumped because we did not have equations for more than three alleles under the egg-dumping model. Those probabilities were also combined using equation (5) to give overall probabilities for each type of exclusion.

It is important to realize that the overall probability of detection calculated from the data in Table 3 with equation (5) is not equivalent to the sum of the probabilities of each type of exclusion. Equation (5) calculates the probability of detecting an exclusion at one *or more* loci. Similarly, the overall probability of each type of exclusion is the probability of detecting that type of exclusion at one or more loci; additional exclusions of that type at other loci are ignored. However, in rare instances offspring will have genotypes that, for example, exclude the putative father at one locus and exclude one of the parents (ambiguous exclusion) at another locus. In analyses of the distribution of types of exclusion, these individuals are counted under both types of exclusion. This creates no problem with independence as long as the loci segregate independently.

To investigate which type of stray genes (EPCs or egg-dumping) occurred in indigo buntings we

Table 3. Allele frequencies, relative mobilities, standard deviations, and probabilities of detection for nine polymorphic loci found in indigo bunting muscle (Westneat 1987b)

Enzyme	Relative mobility	Allele frequency	*N*	Probability of detection (d_i)
a-GPD	100	0.973±0.011	220	0.0209
	60	0.018±0.009		
	135	0.009±0.006		
6-PGD	100	0.926±0.017	230	0.0732[a]
	60	0.035±0.012		
	148	0.022±0.010		
	120	0.013±0.007		
	208	0.004±0.004		
GOT	100	0.924±0.017	236	0.0706
	27	0.059±0.008		
	200	0.017±0.008		
PGM	100	0.979±0.009	234	0.0208
	70	0.013±0.007		
	144	0.004±0.004		
	217	0.004±0.004		
PGI	100	0.911±0.019	236	0.0745
	240	0.089±0.019		
IDH	100	0.976±0.011	206	0.0229
	185	0.024±0.011		
MPI	100	0.949±0.018	156	0.0488
	105	0.038±0.015		
	94	0.013±0.009		
PEPTB	100	0.897±0.020	232	0.0983
	115	0.052±0.015		
	79	0.047±0.014		
	90	0.004±0.004		
PEPTC	100	0.931±0.017	234	0.0655
	114	0.051±0.014		
	90	0.009±0.006		
	74	0.009±0.006		

Overall probability of detection $= 1 - \prod_{i=1}^{n} (1 - d_i) = 0.401$

[a] The two rarest alleles were lumped; this did not affect the number of exclusions

Table 4. Partitioning of the probability of detection into probabilities for each type of exclusion (based on the data in Table 3 and from Westneat 1987b)

Type of exclusion	Both male and female	Male only	Female only	Ambiguous exclusion
EPC model				
Absolute frequency				
1983	0.000	0.063	0.000	0.331
1984	0.000	0.066	0.000	0.357
Expected proportion of types				
1983	0.000	0.159	0.000	0.841
1984	0.000	0.156	0.000	0.844
Egg-dumping model				
Absolute frequency				
1983	0.016	0.036	0.036	0.352
1984	0.013	0.037	0.037	0.346
Expected proportion of types				
1983	0.036	0.081	0.081	0.802
1984	0.030	0.086	0.086	0.798

Table 5. Comparison of observed types of exclusion with predictions from two models of stray genes (EPC and Egg-dumping). These data are from Westneat (1987b). The expected distributions are based on the frequencies in Table 4 and the total number of observed exclusions

Types of exclusion	Both male and female	Male only	Female only	Ambiguous exclusion
Observed distribution				
1983	0	4	0	6
1984	0	6	0	23
Total	0	10	0	29
Expected distribution: EPC model				
1983	0.0	1.6	0.0	8.4
1984	0.0	4.5	0.0	24.5
Total	0.0	6.1	0.0	32.9
Expected distribution: Egg-dumping model				
1983	0.4	0.8	0.8	8.0
1984	0.9	2.5	2.5	23.1
Total	1.3	3.3	3.3	31.1

compared the observed distribution of types of exclusions with those predicted by both of the models. To achieve a reasonable sample size for this comparison, we have added data from 1983 to that of 1984. Comparison of offspring genotypes with those of putative parents revealed a total of 10 cases of exclusion ($N=97$ offspring analyzed) in 1983 and 27 cases of exclusion in 1984 ($N=160$ offspring) (Westneat 1987b). Only two types of exclusion were observed; 8 young had genotypes that excluded the putative father, 27 young had genotypes that excluded one or the other of the putative parents (ambiguous exclusion), and two young had genotypes that excluded the male at one locus, and excluded one or the other of the putative parents at another locus. Thus a total of 39 exclusions were observed.

The observed distribution of types of exclusions is shown in Table 5. To calculate the expected distributions under each model, we calculated the expected proportion of types of exclusion from each model by dividing the probabilities of detection for each type by the total probability of detection (see Table 4). Finally, the 39 exclusions were partitioned into the expected distributions using the expected proportions of each type of exclusion under each model (Table 5).

Comparison of the observed distribution to the expected suggests that the observed exclusions came from extra-pair fertilizations. No female exclusions of either type (female only or both male and female) were observed. In fact the observed distribution of exclusions is significantly different from that expected from the egg-dumping model ($G=12.3$, df$=3$, $P<0.01$), and not significantly different from the expectation of the EPC model ($G=1.98$, df$=1$, $P>0.05$; Sokal and Rohlf 1969; see Table 5). Nest checks made during egg-laying confirmed that egg-dumping is at least a rare event in this population (Westneat 1987b, unpublished data).

Estimating the frequency of stray genes

If egg-dumping is very rare in buntings, then all the parental exclusions in 1984 (16.9% of all offspring sampled) were the result of extra-pair fertilizations. However, since the probability of detection under the EPC model was 0.401 in 1984, the observed 27 cases (offspring with genotypes excluding putative parents) in that year should only be 40% of the actual number of cases. A simple back calculation, dividing the observed rate by the probability of detection, provides an estimate of the actual rate of extra-pair fertilizations (0.421 of all young).

There is, however, some sampling error asso-

ciated with this estimate. First, the observed frequency of cases (0.169) will be distributed as a binomial. This observed rate will therefore have a standard deviation of $\sqrt{(0.169)(0.831)/160}$ (Hays 1981), which is 0.029.

Second, since the probability of detection depends on the frequencies of alleles, then the error in each d_i will be a function of the error in the allele frequencies. Allele frequencies are also distributed as binomials, so again $\sigma = \sqrt{pq/N}$. In the cases of loci with three and four alleles, the standard deviation for each allele was calculated as a binomial with all the other allele frequencies lumped. To calculate the standard deviation, s_i on d_i, we used values for the allele frequencies within their standard deviation that gave the most extreme deviation in d_i.

The error in the overall probability of detection (D) can be calculated for multiple loci by a method developed by Dana Quade of the Department of Biostatistics at the University of North Carolina, Chapel Hill. If loci are independent, and the number of loci is not too small (less than 8 to 10), then the confidence interval is

$$1-\left[\prod_{i=1}^{n}(1-d_i)\,e^{\pm z\sqrt{\sum_{i=1}^{n} s_i^2/(1-d_i)^2}}\right] \qquad (8)$$

where z is the normal critical value, d_i is the probability of detection for the ith locus, and s_i^2 is the variance in d_i for the ith locus (see Appendix for the derivation).

Because the derivation of equation (8) is based on a logarithmic function, the confidence interval is not symmetrical. For the allele frequencies in Table 3 the probability of detection of 0.401 has a 95%-confidence interval of 0.343 to 0.454.

To calculate the confidence interval for the estimated rate of extra-pair fertilizations, we combined the confidence interval on the observed frequency of extra-pair fertilizations with the confidence interval on the probability of detection (D). This gives an interval larger than 95%, since sampling errors in both the probability of detection and the observed rate are not expected to vary in the same direction to the same extent. For the bunting data of 1984, the 95+%-confidence interval on the estimated rate of extra-pair fertilizations is 0.247 to 0.659.

Discussion

Estimates of the probability of detecting a case of stray genes allow a much more quantitative use of electrophoretic data. For indigo buntings this procedure has allowed the calculation of rates of extra-pair fertilizations and the elimination of egg-dumping as a significant behavioral cause of stray genes.

A critical point of discussion is that the assumptions of our models might not apply in all cases. First, all alleles must fit the Hardy-Weinberg distribution. We assumed the one locus (PGI) in the buntings in 1984 did not fit the Hardy-Weinberg distribution as a result of sampling error, because (1) that locus was very close to Hardy-Weinberg equilibrium in 1983, (2) out of 17 statistical tests one is fairly likely to be significant by chance alone, (3) all of the population was not sampled. Some species, especially those with restricted dispersal patterns might not have genotype frequencies in Hardy-Weinberg equilibrium, and some enzymes, such as serum esterases, have notoriously unpredictable banding patterns (McGovern and Tracy 1981; Mumme et al. 1985). A test for deviations from Hardy-Weinberg predictions is therefore prerequisite to any use of these models.

One problem with testing for fit to the Hardy-Weinberg distribution is that one accepts the null hypothesis of no difference. This is dangerous if the power of the statistical test is low. For the bunting data, the power of the G-test used was 60% against the alternative that there was a 30% skew from Hardy-Weinberg expectations in the number of heterozygotes (at a single locus). This skew would result in a 23% difference in the probability of detection for that locus, calculated by using genotype frequencies instead of allele frequencies (see below). However, the overall probability based on nine loci would be in error by only 3%. In addition, the direction of this difference depends on the direction of the skew in the number of heterozygotes. If one concluded erroneously that a locus fit Hardy-Weinberg prediction when in actuality there was a surplus of heterozygotes, then one would be overestimating the probability of detection. If several loci are used to calculate the probability of detection, this becomes less of a concern, unless most of the loci are skewed in the same direction.

Our models can be modified to correct for deviations from Hardy-Weinberg expectations by using genotype frequencies rather than allele frequencies. Instead of plugging allele frequencies into the general equations given in this paper, one could use the observed frequencies of genotypes to calculate the probability of each mating arrangement. Wrege and Emlen (1987) used this procedure in their study of parentage in bee-eaters. The use of

observed genotype frequencies does have some drawbacks. First, the equations for calculating the probability of detection become more cumbersome, especially with 3 or 4-allele loci. This occurs primarily because using genotypes increases the number of variables (p and q replaced with p^2, $2pq$, and q^2). The use of genotype frequencies also increases the error on the probability of detection (D), since by switching to genotypes the sample size is reduced by a half. Unless there is some reason to believe several loci in the population do not fit Hardy-Weinberg expectations, we advocate the use of allele frequencies instead of genotype frequencies.

Combining probabilities of detection from different loci depends on using independently segregating loci. In most cases, loci are unlikely to be linked since they are assayed at random with respect to their locations on chromosomes. Linkage would reduce the probability of detection. Equations for calculating the probability of detection that take linkage into account have been developed for human parental exclusion (Smouse and Adams 1983) and could be adapted for addressing issues in sociobiology.

In our models we also assume that alleles are codominant. For most soluble proteins, codominance is the most common situation (Manwell and Baker 1970). But some genetic markers, such as plumage polymorphisms, do not have codominant alleles. For these cases, models designed to take dominance into account could be adapted from the literature on human parental exclusion (Neel and Schull 1954; MacCluer and Schull 1963).

Finally, we have assumed that matings occur randomly with respect to genotype. In some species, EPCs or egg dumping are likely to involve relatives. Mumme et al. (1985) found this problem in their analysis of acorn woodpeckers. MacCluer and Schull (1963) and Salmon and Brocteur (1978) have derived models to calculate probabilities of detecting close-relative matings in humans.

An interesting ramification of equations to calculate the probability of detection was the use of these equations to separate the behavioral causes of stray genes. However, separation of the behavioral causes of stray genes might be more difficult in other species in which both egg-dumping and EPCs are known to occur (Frederick 1987; Wrege and Emlen 1986; 1987). Our models could be applied to such cases by fitting the observed distribution of types of exclusions to a weighted mixture of the EPC and egg-dumping models using a maximum likelihood procedure (Westneat, unpublished data). For example if 85% weighting of the EPC model and 15% weighting of the egg-dumping model provided the best fit to the observed distribution of types of exclusions, one could use these weightings as estimates of the relative proportions of the observed exclusions resulting from these tactics. Use of the appropriate probability of detection would give an estimate of the frequency of the number of young resulting from each reproductive option.

Table 6. Probabilities of detection for some published analyses of soluble enzymes found in the blood and muscle of passerine birds. Species with the largest number of individuals genotyped were selected from the references. Probabilities of detection were calculated under the EPC model

Species	Number of polymorphic loci	Probability of detection
Catharus ustulatus[a]	4	0.260
Catharus guttatus[a]	6	0.446
Dendroica coronata[b]	11	0.756
Seirus noveboracensis[b]	13	0.915
Passerella iliaca[c]	11	0.441
Junco hyemalis[c]	7	0.457

[a] Avise et al. (1980)
[b] Barrowclough and Corbin (1978)
[c] Zink (1982)

The use of such best fit procedures to estimate the relative proportions of stray genes caused by EPCs or egg-dumping must be viewed with some cuation. Most parental exclusions are ambiguous (see Tables 4 and 5). If a small number of exclusions are observed, the error on the best fit will probably be large. This means that researchers should consider both the quality of the estimates they wish to obtain and the sampling effort needed when contemplating a study of parentage in species with both EPCs and egg-dumping.

A related concern about the use of genetic markers for parental exclusion is that the number of polymorphic enzymes obtained is often so small that the probability of detection is slight (Barrowclough et al. 1985; Mumme et al. 1985). For instance, several researchers have found very few polymorphic enzymes in avian blood, and studies using this tissue alone have led to inconclusive results Frederick 1985; Mumme et al. 1985). However, the published data show much higher isozyme variability in avian muscle tissue (Table 6), and the development of relatively benign biopsies for muscle tissue (Baker 1981; Westneat 1986; Westneat et al. 1986) and feather pulp (Marsden and May 1984) will probably allow the use of these tissues in a variety of birds (but see Frederick 1986). Thus through the use of more productive tissues and

the estimation of error we present here, electrophoretic studies of parental exclusion are probably feasible in many more species than has been previously thought.

At the same time, we second the cautions of Mumme et al. (1985) in recommending a pilot study before any full-scale investigation of parentage is actually undertaken. Preliminary sampling of 20–30 individuals should give an idea of the feasibility of obtaining the desired tissue and also working estimates of the enzyme variability and of the probability of detection. Behavior suggesting stray genes, such as extra-pair consortships and copulations or dumped eggs, should be quantified to obtain preliminary estimates of the frequencies and types of stray genes in the population. However, observational methods can underestimate the actual rates of stray genes (Westneat 1987a, b). If a sufficient sample size (200 young) and a reasonable probability of detection (0.3–0.5) can be obtained, electrophoresis will provide important information no matter how few cases of exclusion are detected.

Finally, our models have been designed with birds in mind. However, other organisms can satisfy the assumptions of our models. The EPC model applies in any situation in which a particular male's parentage is in question. In mammals, males often defend harems of females (Le Boeuf 1974; Clutton-Brock et al. 1982). In many of these situations, subordinate males appear to gain some copulations with females (Duvall et al. 1976; Curie-Cohen et al. 1983). Paternal exclusion in these situations might allow more precise estimation of the reproductive success gained by both dominant and subordinate males. The egg-dumping model similarly would apply to species, such as fish (Baylis 1981) or insects (Tallamy 1985), in which individuals sometimes care for eggs laid by another individual. It is our hope that the methods we have addressed in this paper will provide some additional tools for research on alternative reproductive tactics in a variety of species.

Acknowledgments. We are indebted to Dr. Dana Quade for deriving the formula for confidence intervals on the overall probability of detection and for reviewing the manuscript. Valuable suggestions and criticisms of early drafts of this paper were contributed by Robert Payne, Helmut Mueller, Nelson Hairston, Alan Stiven, Charles Aquadro, Patricia Gowaty, Walter Piper, Chris Verlinden, Ted Stevens, and six anonymous reviewers. Westneat's study was supported by grants from the National Science Foundation (BSR 8501075), Sigma Xi, The Frank M. Chapman Memorial Fund, the Department of Biology of the University of North Carolina, Chapel Hill, and by NSF grants to Robert Payne (BNS 8102404, BSR 8317810). We also thank Steve Emlen for discussion of the methods presented here. He, Peter Wrege, and Christopher Davies have independently and simultaneously developed a similar procedure to analyze parental exclusions in bee-eaters (Wrege and Emlen 1987).

Appendix

Dana Quade of the Department of Biostatistics at the University of North Carolina, Chapel Hill, derived the following method for calculating the confidence interval on the overall probability of detection (D):

The overall probability of detection (Δ) is a product of the probabilities of detection (δ_i) for i independent loci;

$$1-\prod_{i=1}^{n}(1-\delta_i) \tag{1}$$

Since the expectation of δ_i is d_i and the variance is s_i^2, the confidence interval on Δ will be a function of the variance for each $\delta_i(s_i^2)$.

Suppose $t_i=\log(1-d_i)$. The expectation of t_i is therefore approximately

$$\log(1-\delta_i) \tag{2}$$

and the estimated variance (from a Taylor expansion) of t_i is

$$s_i^2/(1-d_i)^2 \tag{3}$$

The expectation of the sum of t_i, $\sum_{i=1}^{n} t_i$, is

$$\sum_{i=1}^{n}\log(1-d_i) \tag{4}$$

with a variance of

$$\sum_{i=1}^{n} s_i^2/(1-d_i)^2 \tag{5}$$

If the number of independent d_i is not too small (no less than 8–10), then $\sum_{i=1}^{n} t_i$ should be approximately normal. Thus, the confidence interval on $\sum_{i=1}^{n}\log(1-\delta_i)$ is

$$\sum_{i=1}^{n}\log(1-d_i)\pm z\sqrt{\sum_{i=1}^{n} s_i^2/(1-d_i)^2} \tag{6}$$

and the confidence interval on $1-\prod_{i=1}^{n}(1-\delta_i)$ is

$$1-\left[\prod_{i=1}^{n}(1-d_i)\,e^{\pm z\sqrt{\sum_{i=1}^{n} s_i^2/(1-d_i)^2}}\right] \tag{7}$$

References

Avise JC, Patton JC, Aquadro CF (1980) Evolutionary genetics of birds I. Relationships among North American thrushes and allies. Auk 97:135–147

Baker MC (1981) A muscle biopsy procedure for use in electrophoretic studies of birds. Auk 98:392–393

Barrowclough GF, Corbin KW (1978) Genetic differentiation in the Parulidae. Auk 95:691–702

Barrowclough GF, Johnson NK, Zink RM (1985) On the nature

of genic variation in birds. In: Johnston RF (ed) Current ornithology, Vol 2. Plenum Press, New York, pp 135–154
Baylis JR (1981) The evolution of parental care in fishes, with reference to Darwin's rule of male sexual selection. Env Biol Fish 6(2):223–251
Birdsall DA, Nash D (1973) Occurrence of successful multiple insemination of females in natural populations of Deer Mice *Peromyscus maniculatus*. Evolution 27:106–110
Birkhead TR, Johnson SD, Nettleship DN (1985) Extra-pair matings and mate guarding in the Common Murre *Uria aalge*. Anim Behav 33:608–619
Brown CR (1984) Laying eggs in a neighbor's nest: benefit and cost of colonial nesting in Swallows. Science 224:518–519
Burns JT, Cheng KM, McKinney F (1980) Forced copulation in captive Mallards. I. Fertilization of eggs. Auk 97:875–879
Chakraborty RM, Shaw M, Schull WJ (1974) Exclusion of paternity: the current state of the art. Am J Hum Genet 26:477–488
Chakravarti A, Li CC (1983) The effect of linkage on paternity calculations. In: Walker RH (ed) Inclusion probabilities in parentage testing. American Association of Blood Banks, Arlington, 439 p
Cheng KM, Burns JT, McKinney F (1983) Forced copulation in captive Mallards III: Sperm competition. Auk 100:302–310
Clutton-Brock TH, Guiness FE, Albon SD (1982) Red Deer – behavior and ecology of two sexes. Univ Chicago Press, Chicago
Cox CR, Le Boeuf BJ (1977) Female incitation of male competition: A mechanism in sexual selection. Am Nat 111:317–335
Curie-Cohen M, Yoshihara D, Luttrell L, Benforado K, MacCluer JW, Stone WH (1983) The effects of dominance on mating behavior and paternity in a captive troop of Rhesus Monkeys (*Macaca mulatta*). Am J Primat 5:127–138
Davies NB (1985) Cooperation and conflict among dunnocks, (*Prunella modularis*) in a variable mating system. Anim Behav 33:628–648
Dunbar RM (1985) Reproductive Decisions. Cambridge University Press, Cambridge
Duvall SW, Bernstein IS, Gordon TP (1976) Paternity and status in a Rhesus monkey group. J Reprod Fert 47:25–31
Emlen ST, Wrege PH (1986) Forced copulation and intra-specific parasitism: Two costs of social living in the White-fronted Bee-eater. Ethology 71:2–29
Fisher HA (1971) The Laysan Albatross: its incubation, hatching and associated behaviors. Living Bird 10:19–78
Fitch MA, Shugart GW (1984) Requirements for a mixed reproductive strategy in avian species. Am Nat 124(1):116–126
Foltz DW (1981) Genetic evidence for long term monogamy in a small rodent, *Peromyscus polionotus*. Am Nat 117(5):665–675
Foltz DW, Hoogland JL (1981) Analysis of the mating system in the Black-tailed Prarie Dog (*Cynomys ludovidianus*) by likelihood of paternity. J Mamm 62(4):706–712
Ford NL (1983) Variation in mate fidelity in monogamous birds. In: Johnston RF (ed) Current ornithology (I). Plenum Press, New York, pp 329–356
Frederick PC (1985) Mating strategies of White Ibis (*Eudocimus albus*). PhD diss, Univ of North Carolina, Chapel Hill
Frederick PC (1986) Parental desertion of nestlings by white ibis (*Eudocimus albus*) in response to muscle biopsy. J Field Ornithol 57:168–170
Frederick PC (1987) Extra-pair copulations in the mating system of White Ibis (*Eudocimus albus*). Behaviour 100:170–201
Gavin TA, Bollinger EK (1985) Multiple paternity in a territorial passerine: the Bobolink. Auk 102:550–555
Gladstone DE (1979) Promiscuity in monogamous colonial birds. Am Nat 114:545–577
Gowaty PA, Karlin AA (1984) Multiple maternity and paternity in single broods of apparently monogamous Eastern Bluebirds (*Sialia sialis*). Behav Ecol Sociobiol 15:91–95
Hanken J, Sherman PW (1981) Multiple paternity in Beldings Ground Squirrel litters. Science 212:351–353
Hays WL (1981) Statistics, 3rd ed. Holt, Rinehart, and Winston, New York
Joste NE, Ligon JD, Stacey PB (1985) Shared paternity in the Acorn Woodpecker (*Melanerpes formicivorus*). Behav Ecol Sociobiol 17:39–41
Le Boeuf BJ (1974) Male-male competition and reproductive success in Elephant Seals. Am Zool 14:163–176
MacCluer JW, Schull WJ (1963) On the estimation of the frequency of non-paternity. Am J Hum Genet 15:191–202
Manwell C, Baker CF (1970) Molecular biology and the origin species: Heterosis, protein polymorphism, and animal breeding. Univ Washington Press, Seattle
Marsden JE, May B (1984) Feather pulp: a non-destructive sampling technique for electrophoretic studies of birds. Auk 101:173–174
McGovern MG, Tracy CR (1981) Phenotypic variation in electromorphs previously considered to be genetic markers in *Microtus ochrogaster*. Oecologia (Berlin) 51:276–280
McKinney F, Cheng KM, Bruggers DJ (1984) Sperm competition in apparently monogamous birds. In: Smith RL (ed) Sperm competition and the evolution of animal mating systems. Academic Press, New York, pp 523–545
McKracken GF, Bradbury JW (1977) Paternity and genetic heterogeneity in the polygynous bat, *Phyllostomus hastatus*. Science 198:303–306
Mock DG (1983) On the study of avian mating systems. In: Brush AH, Clark GA (eds) Perspectives in ornithology. Cambridge University Press, New York, pp 55–85
Mumme RL, Koenig WD, Zink RM, Marten JM (1985) An analysis of genetic variation and parentage in a California population of Acorn Woodpeckers. Auk 102:305–312
Neel JV, Schull WJ (1954) Human heredity. Univ Chicago Press, Chicago
Oring LW (1982) Avian Mating Systems. In: Farner DS, King JR, Parkes KC (eds) Avian biology, vol 6. Academic Press, New York, pp 1–92
Pierotti RJ (1980) Spite and altruism in gulls. Am Nat 115:290–300
Rasmussen DI (1964) Blood group polymorphism and inbreeding in natural populations of the Deer Mouse *Peromyscus maniculatus*. Evolution 18:219–229
Salmon DB, Brocteur J (1978) Probability of paternity exclusion where relatives are involved. Am J Hum Genet 30:65–75
Sherman PW (1981) Electrophoresis and avian genealogical analyses. Auk 98:419–422
Silver HS (1978) Paternity testing: A seminar presented by the Committee on Technical Workshops. American Association of Blood Banks, Washington
Silver HS (1982) Probability of inclusion in paternity testing: A technical workshop. American Association of Blood Banks, Arlington
Smith RL (1984) Sperm competition and the evolution of animal mating systems. Academic Press, New York
Smouse PE, Adams JA (1983) The use of closely linked genetic markers for paternity analysis, with special consideration of the HLA complex. In: Walker RH (ed) Inclusion probabilities in parentage testing. American Association of Blood Banks, Airlie
Sokal RR, Rohlf FJ (1969) Biometry. Freeman, San Francisco

Tallamy DW (1985) "Egg-dumping" in lace bugs (*Gargaphia solami*, Hemiptera: Tingidae). Behav Ecol Sociobiol 17:357–362

Van Tienhoven A (1983) Reproductive physiology of vertebrates, 3rd ed. Cornell University Press, Ithaca

Werschkul DF (1982) Nesting ecology of the Little Blue Heron: promiscuous behavior. Condor 84:381–384

Westneat DF (1987a) Extra-pair copulations in a predominantly monogamous bird: observations of behaviour. Anim Behav 35:865–876

Westneat DF (1987b) Extra-pair fertilizations in a predominantly monogamous bird: genetic evidence. Anim Behav 35:877–886

Westneat DF (1986) The effects of muscle biopsy on survival and condition in white-throated sparrows. Wilson Bull 98:280–285

Westneat DF, Payne RB, Doehlert SM (1986) Effects of muscle biopsy on survival and breeding success in Indigo Buntings. Condor 88:220–227

Wiener AS, Lederer M, Polayes SH (1930) Studies in isohemoglutination, IV, on the changes of providing non-paternity with special reference for blood groups. J Immunol 19:259–282

Wrege PH, Emlen ST (1987) Biochemical determination of parental uncertainty in white-fronted bee-eaters. Behav Ecol Sociobiol 20

Wyllie I (1981) The Cuckoo. Universe, New York

Yom-Tov Y (1980) Intraspecific nest parasitism in birds. Biol Rev 55:93–108

Zink RM (1982) Patterns of genic and morphologic variation among sparrows in the genera *Zonotrichia*, *Melospiza*, *Junco*, and Passerella. Auk 99:632–638