Theor Appl Genet (1987) 75:157–164

Progeny testing for recessive genes: procedures and interpretations

H. Tanida and W. Hohenboken

Department of Animal Science, Oregon State University, Corvallis, OR 97331, USA

Received April 3, 1987; Accepted June 16, 1987
Communicated by L. D. Van Vleck

Summary. Methods for calculating the probability of detecting a carrier of a recessive gene by utilizing matings among related individuals are presented for single and litter bearing species. The confidence level for detection of heterozygosity depends upon: (1) the genetic relationship between mates, (2) the number of mates per male and the number of offspring per mate, (3) whether an estimate of recessive gene frequency before selection is available and (4) the magnitude of that frequency. Methods of computing probability of heterozygosity vs homozygosity utilizing Bayes theorem also are presented. In the conventional progeny test method, a sire initially is assumed heterozygous before calculations are made, but no prior information concerning his probable genotype is utilized. In the method using Bayes theorem, prior sources of information from relatives or from estimates of population allele frequency are utilized. This method gives the exact probability that a sire is not a carrier, given prior information and that he produces all normal offspring. These methods could be used in any sexually reproducing species to identify not only detrimental genes but beneficial genes as well.

Key words: Progeny testing – Recessive genes – Bayes Theorem – Carrier detection

Introduction

Animals phenotypically abnormal due to homozygosity for a recessive gene can readily be detected and eliminated from a population. When recessive homozygosity causes death prior to reproduction or inability to reproduce, recessive gene frequency is automatically reduced in subsequent generations. When, however, recessive allele frequency is low, the gene occurs predominantly in heterozygotes; and there are few homozygotes upon which selection can act. Further reduction in recessive gene frequency becomes very slow.

A solution to this problem is to carry out test matings for heterozygote detection. An effective procedure is to mate a suspected carrier male to known homozygous recessive or known carrier females. When one or more homozygous recessive offspring are produced, it proves that the tested male was, in fact, a carrier. When no homozygous recessive offspring are found, it can be stated, with a confidence level which is a function of the probability of producing a homzygous recessive offspring in a single mating and of the number of normal offspring produced by the matings, that the tested male probably is not a carrier.

Females of the appropriate, known genotype, however, often are unavailable, and for many conditions homozygous recessive females cannot reproduce. Even when genotypes of tester females are unknown, detecting a recessive gene carrier is possible utilizing matings among related individuals. Although confidence levels for carrier detection using females of known genotype are easily calculated, the methods of calculating confidence levels from results of matings among related individuals have not been reported in a complete manner.

When prior information on genotypes of a male's relatives and/or allele frequencies at the locus of interest are available, a method utilizing Bayes theorem, which incorporates the prior information, can be used in place of the conventional method. Bayes theorem is used routinely in genetic counseling concerning probabilities of expression of recessive genes in humans.

The objectives of this paper are: (1) to present explicit procedures for calculating confidence levels from a carrier detection test by sire-daughter, full sib and half sib matings, as well as by combinations of these and other mating types and (2) to present methods using Bayes theorem.

Assumptions

Assumptions common to all following examples are: (1) we are dealing with an autosomal locus with two alleles, A and a which exist at frequencies of p and q, respectively. (2) Dominance is complete, i.e. (AA = Aa) ≠ aa. (3) Homozygous recessive offspring (aa) cannot reproduce. (4) Known heterozygous females are not available.

The conventional progeny test method of carrier detection

Case 1: carrier testing by mating a male to females drawn at random from the population

This case has been presented by Van Vleck et al. (1987) for single bearing species, but it will be explained here because matings among related individuals proceed from this basic mating. The required initial hypothesis is that the male is a carrier (Aa). The probability that he produces phenotypically normal offspring when mated to females drawn at random from the population is then computed, conditional upon the assumption of his heterozygosity. Let p and q be frequencies of dominant and recessive alleles (A and a) in the previous generation. Recessive homozygotes cannot reproduce, and so frequencies of AA and Aa in potential mates in the current generation must be recalculated and sum to one. Expected proportions of AA and Aa potential mates before selection (p^2 and $2pq$, respectively) are summed and each proportion is divided by that sum. This gives genotypic frequencies of $\frac{1-q}{1+q}$ for AA and $\frac{2q}{1+q}$ for Aa in potential mates.

When a mate of the male is AA (with probability of $\frac{1-q}{1+q}$), offspring always will be phenotypically normal (AA or Aa). The joint probability of picking an AA female as mate and producing a normal offspring is the product of the probabilities or $\left(\frac{1-q}{1+q}\right)(1)$. In litter bearing species (or a single bearing species with a female producing n offspring in her lifetime), the probability that all n offspring will be normal is $\left(\frac{1-q}{1+q}\right)(1)^n$. The power series is appropriate because each subsequent offspring is independent of each previous one. In this case, $1^n = 1$ and the joint probability always reduces to $\frac{1-q}{1+q}$.

When the mate is Aa (with probability $\frac{2q}{1+q}$), the probability of producing a phenotypically normal offspring (AA or Aa) is 3/4. The joint probability of picking an Aa female and producing one normal offspring is $\left(\frac{2q}{1+q}\right)\left(\frac{3}{4}\right)$, and the probability of producing n normal and no abnormal offspring is $\left(\frac{2q}{1+q}\right)\left(\frac{3}{4}\right)^n$.

Since both events, picking an AA female or an Aa female as mate, are mutually exclusive, the overall probability of producing n normal offspring by mating a male to a single female drawn at random from the population is the sum $\left(\frac{1-q}{1+q}\right)+\left(\frac{2q}{1+q}\right)\left(\frac{3}{4}\right)^n$.

The male could be mated with any number (N) of females, and the females could produce variable numbers (n_i) of offspring. The probability of producing all normal offspring $\left(\sum_{i=1}^{N} n_i\right)$ by mating the male to N females drawn at random, if in fact he is a carrier, is

$$\left(\frac{1-q}{1+q}\right)+\left(\frac{2q}{1+q}\right)\left(\frac{3}{4}\right)^{n_1}\times\left(\frac{1-q}{1+q}\right)+\left(\frac{2q}{1+q}\right)\left(\frac{3}{4}\right)^{n_2}\times\ldots$$

$$\times\left(\frac{1-q}{1+q}\right)+\left(\frac{2q}{1+q}\right)\left(\frac{3}{4}\right)^{n_N}$$

$$=\prod_{i=1}^{N}\left[\left(\frac{1-q}{1+q}\right)+\left(\frac{2q}{1+q}\right)\left(\frac{3}{4}\right)^{n_i}\right].$$

When only one offspring is produced per female ($n_i = 1$) the formula reduces to

$$\prod_{i=1}^{N}\left[\left(\frac{1-q}{1+q}\right)+\left(\frac{2q}{1+q}\right)\left(\frac{3}{4}\right)^{1}\right]$$

$$=\left[\left(\frac{1-q}{1+q}\right)+\left(\frac{2q}{1+q}\right)\left(\frac{3}{4}\right)\right]^N=\left(\frac{2+q}{2(1+q)}\right)^N.$$

As one offspring per female is a special case of n_i offspring per female, the formula will first be explained for the monotosous case. The confidence of detection is 1 minus the probability of obtaining all normal offspring even if the male actually is a carrier $1-\left(\frac{2+q}{2(1+q)}\right)^N$. The probability of producing all

normal offspring decreases curvilinearly as N increases. If no recessive offspring are produced in spite of a large N, the tested male probably is not a carrier. Thus, confidence level is associated with the conclusion of probable homozygosity. The probability of an incorrect conclusion (the sire is, in fact, a carrier but produced no homozygous recessive offspring in the test mating) is

$$\left(\frac{2+q}{2(1+q)}\right)^N.$$

When there is more than one offspring per female, confidence of detection is

$$1-\prod_{i=1}^{N}\left[\left(\frac{1-q}{1+q}\right)+\left(\frac{2q}{1+q}\right)\left(\frac{3}{4}\right)^{n_i}\right].$$

As the number of offspring per female (n_i) increases, the probability that all n_i offspring are normal decreases and the confidence of detection increases. Also, as the number of mates (N) increases, the probability that all of them produce normal offspring decreases and the confidence of detection increases.

When even one recessive homozygote is found among offspring of a sire, it can be concluded that both parents of the affected offspring are carriers. It is not necessary to conduct further tests nor to compute the probability of detection. When, however, all offspring are normal, the hypothesis that the sire is a carrier may be rejected. Confidence levels would be

$$1-\left(\frac{2+q}{2(1-q)}\right)^N \quad \text{and} \quad 1-\prod_{i=1}^{N}\left[\frac{1-q}{1+q}+\frac{2q}{1+q}\left(\frac{3}{4}\right)^{n_i}\right]$$

for one offspring and n_i offspring per female, respectively.

A large number of normal offspring are necessary to achieve high confidence levels, especially when q is small. For example, for allele frequencies of 0.01 and 0.001, respectively, and one offspring per mate, 604 and 5996 normal offspring would be necessary to achieve a 95% confidence level that a sire was not a carrier. These numbers of offspring are feasible only in unusual circumstances, sires used extensively in artificial insemination, for example. An alternative is to conduct sire-daughter, full sib or half sib matings. For these matings among related individuals, the derivation of the formulae will be shown only in the case of one offspring per female. Complete formulas for the cases of n_i offspring per female are given in Table 1.

Table 1. Confidence levels of rejecting the hypothesis that a tested male is heterozygous for a recessive gene

Mating type	Hypothesis	Single offspring/female	n_i offspring/female
1. Mating to female picked at random	A sire is Aa	$1-\left[\frac{2+q}{2(1+q)}\right]^N$	$1-\prod_{i=1}^{N}\left[\frac{1-q}{1+q}+\frac{2q}{1+q}\left(\frac{3}{4}\right)^{n_i}\right]$
2. Sire-daughter	A sire is Aa	$1-\left[\frac{7+3q}{8+4q}\right]^N$	$1-\prod_{i=1}^{N}\left[\frac{1}{2+q}+\frac{1+q}{2+q}\left(\frac{3}{4}\right)^{n_i}\right]$
3. Combination of 1 and 2	A sire is Aa	$1-\left[\frac{2+q}{2(1+q)}\right]^N\left[\frac{7+3q}{8+4q}\right]^{N'}$	$1-\prod_{i=1}^{N}\left[\frac{1-q}{1+q}+\frac{2q}{1+q}\left(\frac{3}{4}\right)^{n_i}\right]\times\prod_{i=1}^{N'}\left[\frac{1}{2+q}+\frac{1+q}{2+q}\left(\frac{3}{4}\right)^{n_i}\right]$
4. Full sib	A full brother is Aa	$1-\left[\frac{7+3q}{8+4q}\right]^N$	$1-\prod_{i=1}^{N}\left[\frac{1}{2+q}+\frac{1+q}{2+q}\left(\frac{3}{4}\right)^{n_i}\right]$
5. Half sib	A half sib brother is Aa	$1-\left[\frac{15+22q+3q^2}{16+28q+4q^2}\right]^N$	$1-\prod_{i=1}^{N}\left[\frac{3+q}{4+7q+q^2}+\frac{1+6q+q^2}{4+7q+q^2}\left(\frac{3}{4}\right)^{n_i}\right]$
6. Combination of 4 and 5	A half and full brother is Aa	$1-\left[\frac{7+3q}{8+4q}\right]^{N_F}\left[\frac{15+22q+3q^2}{16+28q+4q^2}\right]^{N_H}$	$1-\prod_{i=1}^{NF}\left[\frac{1}{2+q}+\frac{1+q}{2+q}\left(\frac{3}{4}\right)^{NF_i}\right]\times\prod_{i=1}^{N_H}\left[\frac{3+q}{4+7q+q^2}+\frac{1+6q+q^2}{4+7q+q^2}\left(\frac{3}{4}\right)^{NH_i}\right]$

Case 2: carrier testing by mating a sire to his daughters

When offspring produced by mating a sire to randomly chosen females all are normal, the daughters could then be mated back to their sire. By observation, the frequency of aa in daughters is zero, so expected frequencies of AA and Aa in the daughters must be recalculated. These are $\frac{1}{2+q}$ for AA and $\frac{1+q}{2+q}$ for Aa. The probability of producing a normal offspring by mating the sire (hypothesized to be Aa) to a random daughter is calculated in the same manner as in mating the sire to females drawn at random from the population. That is $\frac{1}{2+q}$ (frequency of AA daughters) $\times 1 + \frac{1+q}{2+q}$ (frequency of Aa daughters) $\times \frac{3}{4} = \frac{7+3q}{8+4q}$.

When one or more recessive homozygotes are found among the offspring, it is concluded that both the sire and the particular daughters producing the affected offspring are carriers. Using information from the sire-daughter matings, when all N offspring are normal, the confidence level that the sire is not a carrier is $1 - \left[\frac{7+3q}{8+4q}\right]^N$ and the error rate is $\left[\frac{7+3q}{8+4q}\right]^N$. Some information, however, was already available when the sire produced M normal and no abnormal offspring in the previous generation. Incorporating this information with that from the sire-daughter matings, the confidence becomes $1 - \left\{\left[\frac{2+q}{2(1+q)}\right]^M \times \left[\frac{7+3q}{8+4q}\right]^N\right\}$, which is larger than the confidence level using information from random matings or sire-daughter matings alone.

Case 3: carrier testing by mating a sire to his full sisters

It is not possible from matings among full sibs to test whether the sire of the full sib family is a carrier, since full sib matings could produce homozygous recessive offspring even when the sire was not a carrier. It is possible to test whether a male mated to his own full sibs is a carrier.

If the male to be tested is a carrier, at least one of his parents must have been heterozygous, causing it to contribute A and a alleles to its offspring with equal probability. The other parent could also have been Aa. This other parent would be expected to contribute A and a alleles to its offspring with probabilities of $\frac{1}{1+q}$ and $\frac{q}{1+q}$. Then, the probabilities of the parents producing AA, Aa and aa daughters are $\frac{1}{2} \times \frac{1}{1+q} = \frac{1}{2(1+q)}$, $\frac{1}{2} \times \frac{q}{1+q} + \frac{1}{2} \times \frac{1}{1+q} = \frac{1}{2}$, and $\frac{1}{2} \times \frac{q}{1+q} = \frac{q}{2(1+q)}$, respectively. The aa genotype, however, is assumed not to exist in the family of full sib brothers and sisters. If it had, the original parents would be known carriers, and population allele frequencies would not enter into the calculations. Recalculated genotypic frequencies in the sisters are $\frac{1}{2+q}$ for AA and $\frac{1+q}{2+q}$ for Aa.

The probability of producing a normal offspring by mating the male (hypothesized to be Aa) to a full sister is calculated as before. That is, $\frac{1}{2+q}$ (frequency of AA full sib sisters) $\times 1 + \frac{1+q}{2+q}$ (frequency of Aa full sib sisters) $\times \frac{3}{4} = \frac{7+3q}{8+4q}$. This is equal to the probability that one normal offspring is produced by a single sire-daughter mating (Case 2), when the sire is a carrier. This is reasonable because the genetic relationship of sire and daughter is equal to the genetic relationship of full brother and sister.

The confidence level for rejecting the hypothesis that the male is a carrier when all N offspring are normal is $1 - \left[\frac{7+3q}{8+4q}\right]^N$.

Case 4: carrier testing by mating a male to his paternal half sibs

It also is possible to test whether a male is a carrier by mating him to his paternal half sisters. Once the male is hypothesized to be a carrier, it is known that at least one of his parents must also have been a carrier. A priori, it is equally as probable that the recessive gene was inherited from the sire as from the dam. If it was inherited from the sire, then 50% of paternal half sisters would also be expected to inherit the gene. If, however, the sire to be tested inherited the recessive gene from his dam, then his half sib mates would inherit the recessive gene from their sire only in proportion to its overall frequency in the population. In either case, the dams of half sisters could be heterozygous and would be expected to contribute A and a alleles to their daughters with probabilities $\frac{1}{1+q}$ and $\frac{q}{1+q}$.

When the sire of the paternal half sib is a carrier, he contributes A and a alleles to his offspring with equal probability. When the dam of the half sib brother is a carrier and his sire could also be Aa, this sire would be expected to contribute the A and a allele to his daughters with probabilities of $\frac{1}{1+q}$ and $\frac{q}{1+q}$. As the

events are mutually exclusive and have equal probability, the sire of the paternal half sib family would be expected to contribute the A and a allele to its offspring with probabilities of $\frac{1}{2}\left(\frac{1}{1+q}+\frac{1}{2}\right)=\frac{3+q}{4(1+q)}$ and $\frac{1}{2}\left(\frac{q}{1+q}+\frac{1}{2}\right)=\frac{1+3q}{4(1+q)}$, respectively. The probabilities of producing AA, Aa and aa half sisters are

$$\frac{3+q}{4(1+q)}\times\frac{1}{1+q}=\frac{3+q}{4(1+q)^2},$$

$$\frac{3+q}{4(1+q)}\times\frac{q}{1+q}+\frac{1+3q}{4(1+q)}\times\frac{1}{1+q}=\frac{1+6q+q^2}{4(1+q)^2}$$

and

$$\frac{1+3q}{4(1+q)}\times\frac{q}{1+q}=\frac{q+3q^2}{4(1+q)^2},$$

respectively. Since, however, the aa genotype is assumed not to exist in the family of half brothers and sisters, genotypic frequencies of AA and Aa in half sisters must be recalculated. They equal $\left(\frac{3+q}{4+7q+q^2}\right)$ AA and $\left(\frac{1+6q+q^2}{4+7q+q^2}\right)$ Aa.

The probability of producing a normal offspring by mating the male (hypothesized to be Aa) to a randomly chosen half sister is calculated as before. That is

$$\frac{3+q}{4+7q+q^2}\ \text{(frequency of AA half sib sisters)}\times 1+\frac{1+6q+q^2}{4+7q+q^2}\ \text{(frequency of Aa half sib sisters)}\times\frac{3}{4}=\frac{15+22q+3q^2}{16+28q+4q^2}.$$

The confidence level for rejecting the hypothesis that the half sib brother is a carrier, when all N offspring are normal, is $1-\left(\frac{15+22q+3q^2}{16+28q+4q^2}\right)^N$.

Use of Bayes theorem to compute probabilities of homozygosity vs. heterozygosity

The methods described above give confidence levels of detecting carriers of recessive genes, given the production of no abnormal and variable numbers of normal offspring. With Bayes theorem, it is possible to calculate the exact probability that a sire is a dominant homozygote vs a heterozygote, based upon results of one of the tests already described and consideration of prior quantitative information on the probable genotype of the sire. The general formula for Bayes theorem is:

$$P(A/B)=\frac{P(A)\,P(B/A)}{P(A)\,P(B/A)+P(\bar{A})\,P(B/\bar{A})}.$$

As $\bar{A}$ stands for the nonoccurrence of A, $P(A)+P(\bar{A})=1$. All previously stated assumptions apply here. The case of mating a male to N females drawn at random from the population, with each producing a single offspring, is used as an example, but the concepts described here apply to the other mating schemes as well. Probabilities are defined as:

$P(A)$ = the prior probability that the sire is AA.

$P(\bar{A})$ = the prior probability that the sire is Aa.

If no information is available that the recessive gene is present in close relatives of the male, the prior probabilities that the sire is AA or Aa, respectively, are

$$P(A)=\frac{p^2}{p^2+2pq}=\frac{1-q}{1+q}\quad\text{and}$$

$$P(\bar{A})=\frac{2pq}{p^2+2pq}=\frac{2q}{1+q}.$$

Let $P(B)$ = the probability of producing all normal offspring from a particular type of mating. $P(B/A)$ is the conditional probability of producing all normal offspring, given that the sire is AA. This is equal to 1, since a homozygous sire would transmit A to all its offspring. $P(B/\bar{A})$ is the conditional probability of producing all normal offspring, given that the sire is Aa. It is equal to $\left[\frac{2+q}{2(1+q)}\right]^N$ in this example. (This probability was calculated in the process of calculating the confidence level of detection in the conventional method). The probability that the sire is homozygous AA, given that the matings produce all N normal offspring, is

$$P(A/B)=\frac{\frac{1-q}{1+q}\times 1}{\frac{1-q}{1+q}\times 1+\frac{2q}{1+q}\times\left[\frac{2+q}{2(1+q)}\right]^N}=\frac{\frac{1-q}{1+q}}{\frac{1-q}{1+q}+\frac{2q}{1+q}\left[\frac{2+q}{2(1+q)}\right]^N}.$$

If ancestor or collateral relative information is available, $P(A)$ and $P(\bar{A})$ should be computed from that information. For instance, if both of a sire's parents were carriers, $P(A)=1/3$ and $P(\bar{A})=2/3$. If the sire's grandfather were a known carrier, $P(A)=3/4$ and $P(\bar{A})=1/4$.

When parent-offspring, full sib mating and half sib mating systems were used, $P(B/\bar{A})$ would be

$$\left[\frac{7+3q}{8+4q}\right]^N, \quad \left[\frac{7+3q}{8+4q}\right]^N \quad \text{and} \quad \left[\frac{15+22q+3q^2}{16+28q+4q^2}\right]^N$$

for single bearing species, and

$$\prod_{i=1}^{N}\left[\frac{1}{2+q}+\frac{1+q}{2+q}\left(\frac{3}{4}\right)^{n_i}\right],$$

$$\prod_{i=1}^{N}\left[\frac{1}{2+q}+\frac{1+q}{2+q}\left(\frac{3}{4}\right)^{n_i}\right] \quad \text{and}$$

$$\prod_{i=1}^{N}\left[\frac{3+q}{4+7q+q^2}+\frac{1+6q+q^2}{4+7q+q^2}\left(\frac{3}{4}\right)^{n_i}\right]$$

for litter bearing species.

Discussion

The confidence level from a mating test depends on four factors and the interrelationships among them: (1) the genetic relationship between mates, (2) the number of mates per male and the number of offspring per mate, (3) whether an estimate exists for q, recessive gene frequency before selection and (4) the magnitude of q.

The number of females required to produce normal offspring from each mating system to achieve 90 and 95% confidence levels in litter and single bearing species are given in Table 2. The closer the genetic relationship between a sire and his mates, the fewer mates are required to achieve a given confidence level. Several mating systems can be combined to test a male, as shown earlier. For instance, when sire-daughter matings are utilized, information from mating sires to females to produce daughters should also be available. Incorporating both information sources adds to the confidence of the test $\left(1-\left[\frac{2+q}{2(1+q)}\right]^M \times \left[\frac{7+3q}{8+4q}\right]^N > 1-\left[\frac{7+3q}{8+4q}\right]^N\right)$. As another example, if a male were mated to both full and half sisters to produce offspring, confidence of the test from both information sources would be

$$\left(1-\left[\frac{7+3q}{8+4q}\right]^{N_F} \times \left[\frac{15+22q+3q^2}{16+28q+4q^2}\right]^{N_H}\right).$$

Although formulas for relationships other than sire-daughter, full sib and half sib are not given in this paper, these could be derived in a similar manner. The confidence level will always be increased when information from all available matings is combined, but the increase will sometimes be very small. The more distant the genetic relationship between mates, the less that such information will add to the confidence level. Combinations of information sources are especially useful when the number of offspring produced by each type of mating is limited. Although cases of single offspring per female were used in the examples, the methods would apply equally to cases of more than one offspring per female.

Testing a male by mating to females producing single offspring or to females producing litters requires different total numbers of offspring to achieve any given confidence level (Table 2). For instance, when mates were drawn at random from the population, 464 normal offspring would be needed in a single bearing species to achieve the 90% confidence level, but 1,230 offspring would be needed to achieve the same confidence level in a litter bearing species (assuming $q = 0.01$ and 10 offspring per litter).

Often, the frequency of recessive alleles will not have been estimated. An alternative in such cases is to let $q = 0$ in the formulae. (This does not work in random mating, however, since allele frequency is the only information source from that mating.) This will give a conservative or safer test because a male would have to produce more normal offspring to achieve any given confidence level. It is costly because using $q = 0$ (instead of the true but unknown value) requires larger numbers of mates and/or offspring to achieve the same confidence level.

This conclusion, however, cannot be generalized. It depends on the genetic relationship between a male and his mates. If the genetic relationship between mates is high, such as full sibs or sire and daughter, using $q = 0$ has little effect on the confidence level. For instance, a curve relating confidence (vertical axis) to the number of offspring (horizontal axis, Fig. 1) is not affected when q varies from 0.1 to 0.001 in full sib and sire-daughter

Table 2. The number of mates, all producing only phenotypically normal offspring, required to achieve 90% and 95% confidence levels that a sire is not a carrier, for both single and litter bearing species. Frequency of the recessive allele is assumed to equal 0.01 and litter bearing females are assumed to produce 10 offspring each

	Single bearing species		Litter bearing species	
	90%	95%	90%	95%
Random mating	464	604	123	159
Sire-daughter	18	23	4	5
Full sib	18	23	4	5
Half sib	35	45	9	11

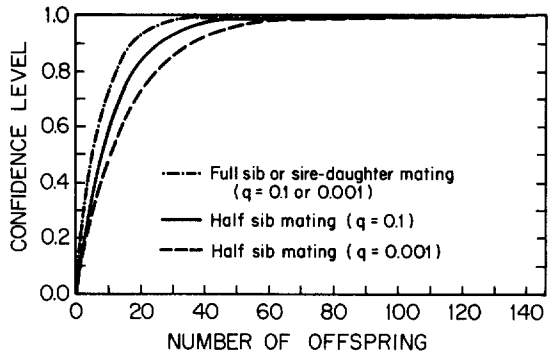

Fig. 1. Confidence levels that a progeny tested male is not heterozygous, based upon results of full sib, sire-daughter or half sib matings and with variable frequencies of a recessive gene

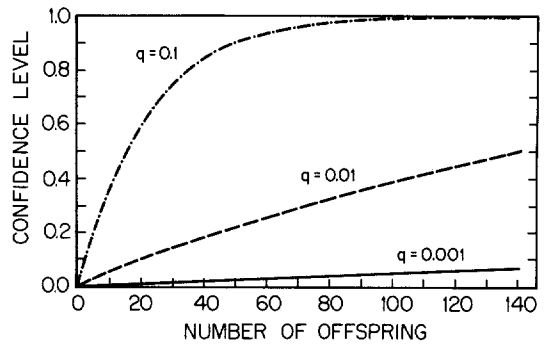

Fig. 2. Confidence levels that a male, mated to females drawn at random from a population, is not heterozygous at three levels of recessive gene frequency

matings. Curves for half sib mating (Fig. 1) and random mating (Fig. 2), however, are mildly and drastically affected by different values of q, and a large N is required to achieve the 90% and 95% confidence levels when q is low (Fig. 2). The differences in sensitivity of the confidence curves to changing values of q correspond to the differences in genetic relationship between a sire and his mates. This is reasonable because a high genetic relationship means that more information is provided by the mates. In random mating, the genetic relationship between the sire and his mates is assumed to be zero, so no information is available from that source. This must be compensated by other sources such as large N and/or high population frequency (q) of the recessive gene.

When known heterozygous (Aa) or homozygous females (aa) are available as testers, mating a male to them is more efficient than mating a male to related individuals. If a female is known to be Aa, only 9 normal offspring are necessary to achieve the 90% confidence level. If a female is aa, only 4 normal offspring are necessary to achieve the same confidence. These methods require that existence of the recessive gene has already been determined; and the test is effective for detecting only that gene. In the method of mating a male to related females, the test is effective for detecting all autosomal recessive genes simultaneously.

The method utilizing relatives, however, has a disadvantage, since it results in inbreeding which likely would result in inbreeding depression for economically important traits.

A few important distinctions should be made between the conventional method and the method using Bayes theorem. In the conventional method, the hypothesis that a sire is a carrier is made before calculation, and no prior information concerning the sire's probable genotype is used. The test result is the confidence of detection or the probability of producing at least one abnormal offspring, assuming that a sire is a carrier. The interpretation of the test result is as follows. If no abnormal offspring are produced, even though the calculated probability of producing at least one abnormal offspring (assuming that the sire is a carrier) is fairly high, the hypothesis that a sire is a carrier is rejected and the alternate hypothesis that a sire is not a carrier is accepted. This is not, however, the probability that the sire is not a carrier.

In the method using Bayes theorem, the hypothesis that a sire is a carrier is not made, and all prior sources of information from relatives or from population allele frequency estimates are utilized. This method gives the exact probability that a sire is not a carrier, given the prior information and that he produces all normal offspring. Since the method based on Bayes theorem utilizes more available information than the conventional method and gives the exact probability that a sire is not a carrier, this method is more powerful. There are, however, a few shortcomings. First, some prior information must be available. Second, prior information must be accurate. If incorrect information is used, the calculated probability from Bayes formula will be biased. Suppose, for example, that a male is mated to full sibs or daughters and that true allele frequency is $q = 0.1$. Producing 5 and 11 normal offspring would result in probability levels exceeding 90% and 95% that a sire was not a carrier. If allele frequency had wrongly been estimated to equal 0.01, however, the calculated probability that the sire was not a carrier would have exceeded 98% with only one normal offspring. At the true allele frequency of 0.1, 18 normal offspring should have been produced to reach the 98% probability level. Thus, the method using Bayes theorem is vulnerable to inaccurate prior information. The conventional method can be less subject to bias from inaccurate estimates of

recessive gene frequency. As shown in Fig. 1, deviating allele frequency from $q = 0.1$ to $q = 0.001$ does not change the required number of normal offspring to reach any given confidence level, when the relationship of the sire to his mates is high.

As a final comment, we have presented and discussed these procedures as if they were effective only for testing males for undesirable genes. This was a simplification and convenience and is not true. In species with high fecundity, chickens and fish for example, the methods could readily be used to test females for heterozygosity. Also the methods could be used in any sexually reproducing species to identify beneficial as well as detrimental genes. With rapid advances in gene transfer technology, we visualize that searches for major genes with beneficial effects will become more common in the future.

Reference

Van Vleck LD, Pollak EJ, Oltenacu EAB (1987) Genetics for the animal sciences. WH Freeman, New York