Theor Appl Genet (1984) 67:317–322

Sex-linked effective population size in control populations, with particular reference to honeybees (*Apis mellifera* L.)

C. Moran
Department of Animal Husbandry, University of Sydney, Sydney, N.S.W. 2006, Australia

Received July 18, 1983
Communicated by J. S. F. Barker

Summary. Sex-linked effective population size (Ne) is derived for a variety of control population structures relevant to normal diploid and/or, more importantly, to haplo-diploid species. For equal sex ratio, it is shown that the control population structure which doubles autosomal effective population size trebles sex-linked effective size. For haplo-diploid species where the number of males exceeds the number of reproductive females, several different control structures are described, which tend to increase effective population size by about 1/3. These would be suitable for stock maintenance of honeybees. Directional selection programmes employing within-family selection would maintain most of the minimum drift/inbreeding properties of these control populations.

Key words: Effective population size – Control population – Sex linkage – Haplo-diploidy – Honeybees

Introduction

Random genetic change is inevitable in populations of finite size. Control populations (Gowe et al. 1959; Latter 1959) are populations set up and maintained in such a way as to minimise this random change.

The rate at which drift/inbreeding[1] occurs is inversely proportional to the effective population size, Ne, a parameter which relates the number of breeding males (M) and breeding females (F) in a population to the number of breeding individuals (N) in an "ideal", random mating, monoecious population with selfing allowed. The extension of this analogy to sex linked loci is strained to say the least, given that sex linkage and monecity are mutually exclusive. Nevertheless, as a means of relating the rate of drift/inbreeding to the number of breeding males and females in a population, the concept is still useful for sex linked loci.

A general expression for the rate of drift/inbreeding for an autosomal locus has been presented by Hill (1972). If no constraints are placed on the contribution of parents along the four autosomal pathways from one generation to the next, the distribution of family size for each pathway is approximately Poisson and the general expression reduces to the familiar formula due to Wright (1931, 1933, 1969). Ollivier (1973) has demonstrated the effects of many different combinations of constraints on family size on autosomal effective population size. For example, if each male parent contributes exactly one son and one daughter to the parents in the next generation and likewise for the female parents, then autosomal effective population size is doubled relative to the situation of no constraints. From the drift/inbreeding point of view, this constitutes the most effective control population.

For sex-linked genes, and for all genes in haplo-diploid organisms, there are a number of important differences from the autosomal situation. First, there are only three pathways for gametes from one generation to the next, namely 1) male parent to daughter, mf, 2) female parent to son, fm, and 3) female parent to daughter, ff. Secondly, along the mf pathway, gene frequency drift can occur only because of sampling between individuals. In the other two pathways (as in all four pathways for diploid autosomal genes), sampling of genes within heterozygotes also contributes to gene frequency drift. An expression for rate of drift at a sex linked locus was developed by Moran and Watterson (1959) for non-overlapping generations but it is not in a suitable form for the purpose of this paper. Pollak (1980) presented the following general formula for sex-linked Ne, which is analogous in form to Hill's (1972) autosomal formula. For the discrete generation

1 If population size is constant from generation to generation (as will be assumed throughout this paper) inbreeding effective population size is equal to variance (drift) effective population size (Crow and Kimura, p 361) and factors which minimise/maximise rate of drift will have the same effect on rate of inbreeding

case it is

$$\frac{1}{Ne} \doteq \frac{1}{9M}\left[2\left(\frac{M}{F}\right)^2 \sigma^2 mf + 1\right] + \frac{1}{9F}\left[\sigma^2 ff + 2\left(\frac{F}{M}\right) cov(fm, ff) + \left(\frac{F}{M}\right)^2 \sigma^2 fm + 1\right] \quad (1.1)$$

where $\sigma^2 ff$ = variance in number of gametes transmitted by females to daughters who become parents in the next generation and $\sigma^2 fm$ and $\sigma^2 mf$ are similarly defined for the other two pathways. cov (fm, ff) = covariance of numbers of male and female progeny produced by females.

In most situations there is no correlation between the numbers of sons and daughters produced by female parents and we can set cov(fm, ff) = 0, as we have done for all subsequent expressions derived from (1.1).

The aim of this paper is to derive formulae for sex-linked effective population size from Pollak's general formula by a progressive application of constraints on family size along the three pathways. Three of these formulae will be compared with those obtained by Page and Laidlaw (1982), who used a different approach. It will be shown that one of their derivations (their formula 5) is erroneous, due to a misunderstanding of constraints which arise in the mf pathway when a system of queen-daughter replacement is employed.

Application of constraints on family size

a) The mf pathway

1. If each male has an equal expectation of producing a daughter, we can approximate the variance of family size by the Poisson variance, which is strictly valid only when the number of parents is not fixed. However this approximation can be shown to give results accurate to second order of 1/M and 1/F. Using this approximation, the mean and variance of daughter number per father is F/M.
2. If M > F, then it is not possible for all male parents to produce a daughter. If those which do produce a daughter are constrained to produce only one, then F/M will produce 1 daughter and (1 – F/M) will produce 0 daughters. There will be a Bernoulli distribution of daughter number, with variance (F/M) (1 – F/M) (Gowe et al. 1959 for autosomal example).
3. If M = F (or nM = F, where n is an integer constant), then all male parents may be constrained to produce exactly 1 daughter (or exactly n daughters). In this case, number of daughters per male parent is constant and variance of family size is zero.

b) The fm pathway

Constraints are applied as for the mf pathway. The Bernoulli variance will apply when F > M, and female parents are restricted to contributing a maximum of one son to the next generation. For zero variance of family size to be possible, either F = M or nF = M, where n is an integer constant.

c) The ff pathway

1. Since the number of females is constant from generation to generation, each female must produce one daughter on average. If no constraints are placed on daughter number and all females have an equal chance of contributing to the next generation, then the distribution of daughter number is approximately Poisson, with mean and variance equal to 1.
2. If each female is constrained to contribute exactly one daughter to the next generation, then clearly the variance of daughter number is zero.

Sex linked effective population size, with specific constraints on family structure

All possible combinations of variances of family size, resulting from applying the abovementioned constraints, have been substituted into Formula 1.1, to produce the 16 values of 1/Ne presented in Table 1. Many of the combinations of constraints on family size presented in Table 1 are unlikely to occur in practice. In particular, cases 11 and 13 are impossible for honeybees because of their mating system and would be extremely difficult to apply even in other organisms where more control could be exerted over the mating system. Those of most significance in reality are marked with a (*).

Results of general significance from Table 1

Case 1 provides a baseline for comparison with the other partly or wholly controlled populations, since no constraints are applied to family size. In this case

$$Ne_{(1)} = \frac{9FM}{4M + 2F} \quad (2.1)$$

which agrees with Wright's (1933) inbreeding derivation and his (1939) variance derivation, and Page and Laidlaw's (1982) formula 3.

By contrast, case 16 represents a rigidly controlled population with zero variance of gamete contribution along the three pathways, and this results in

$$Ne_{(16)} = \frac{9MF}{M + F}. \quad (2.2)$$

Compared with Case 1, setting M = F, the effective population size in this latter case is three times greater.

Wright (1939, 1969) derived the limiting values of effective population size for Formula 2.1 for F much

Table 1. Effect of schemes for controlling gamete choice on sex linked effective population size

Mode of choice of gametes	Case	Distribution and variance of family size			for sex ratio	1/NE
		Path mf	Path fm	Path ff		
Completely random	1*	Poisson [F/M]	Poisson [M/F]	Poisson [1]	Any	$\frac{2}{9M}+\frac{4}{9F}$
Partly random, partly constant number	2*	Poisson [F/M]	Poisson [M/F]	Constant [0]	Any	$\frac{2}{9M}+\frac{3}{9F}$
	3	Poisson [F/M]	Bernoulli [(1–M/F)M/F]	Poisson [1]	F > M	$\frac{2}{9M}+\frac{3}{9F}$
	4	Poisson [F/M]	Bernoulli [(1–M/F)M/F]	Constant [0]	F > M	$\frac{2}{9M}+\frac{2}{9F}$
	5	Constant [0]	Poisson [M/F]	Poisson [1]	F = nM n ≧ 1	$\frac{2}{9M}+\frac{2}{9F}$
	6	Constant [0]	Poisson [M/F]	Constant [0]	F = nM n ≧ 1	$\frac{2}{9M}+\frac{1}{9F}$
	7	Constant [0]	Bernoulli [(1–M/F)M/F]	Poisson [1]	F = nM n > 1	$\frac{2}{9M}+\frac{1}{9F}$
	8*	Constant [0]	Bernoulli [(1–M/F)M/F]	Constant [0]	F = nM n > 1	$\frac{2}{9M}$
	9	Poisson [F/M]	Constant [0]	Poisson [1]	M = nF n ≧ 1	$\frac{1}{9M}+\frac{4}{9F}$
	10*	Poisson [F/M]	Constant [0]	Constant [0]	M = nF n ≧ 1	$\frac{1}{9M}+\frac{3}{9F}$
	11	Bernoulli [(1–F/M)F/M]	Poisson [M/F]	Poisson [1]	M > F	$\frac{4}{9F}$
	12	Bernoulli [(1–F/M)F/M]	Poisson [M/F]	Constant [0]	M > F	$\frac{3}{9F}$
	13	Bernoulli [(1–F/M)F/M]	Constant [0]	Poisson [1]	M = nF n > 1	$\frac{-1}{9M}+\frac{4}{9F}$
	14*	Bernoulli [(1–F/M)F/M]	Constant [0]	Constant [0]	M = nF n > 1	$\frac{-1}{9M}+\frac{3}{9F}$
	15	Constant [0]	Constant [0]	Poisson [1]	M = F	$\frac{1}{9M}+\frac{2}{9F}$
Completely constant	16*	Constant [0]	Constant [0]	Constant [0]	M = F	$\frac{1}{9M}+\frac{1}{9F}$

greater than M and vice versa. In the former case Ne approaches 4.5 M and in the latter Ne approaches 2.25 F. It is noteworthy, but apparently coincidental, that the constraints on family size of case 8 of Table 1 produce an effective population size of 4.5 M, provided only that F = nM and n is an integer greater than 1. Likewise the constraints of case 11 give rise to an effective population size of 2.25 F for any M > F.

Results relevant to sex linked genes in domestic animals

We have already seen, for M = F, that a population structure which doubles autosomal effective population size, trebles sex linked effective population size. However, in populations of domestic animals, the number of breeding females always exceeds the number of breeding males. Application of the Gowe et al. (1959) control population structure, in which each male parent produces exactly one son and exactly F/M(=n) daughters and each female parent produces exactly one daughter, with M random females producing one son and F-M producing no son, in case 8 of Table 1 results in a sex-linked effective size of

$$Ne_{(8)} = 4.5\,M. \tag{2.3}$$

For a female to male sex ratio of n = 2, the control effective population size is twice that of a case 1 population. However, for increasing sex ratio n, the effective size in case 1 asymptotically approaches 4.5 M. The Gowe et al. type of control structure thus has little influence on sex linked effective size when the number

of males is very small relative to the number of females, as for example in artificial insemination programmes.

Results relevant to haplo-diploid organisms

Honeybees, *Apis mellifera,* are the only haplo-diploid organisms in which it may be economically advantageous to establish control or minimum drift populations. Discussion will be restricted to this species, although the results are valid for any haplo-diploid organism.

For male bees, the act of mating is fatal. Queen bees, on the other hand, mate many times, so clearly $M > F$, in contrast to the situation in diploid species of domestic animals. Single inseminations via artificial insemination (A.I.), although possible, are not practicable, and multiple insemination, both in nature and in controlled matings, is the rule. Consequently, exact paternity of daughter queens cannot be controlled and it is difficult to apply constraints to the mf pathway. However, it is easy to apply constraints to the ff and fm pathways and partial constraints may be applied even to the mf pathway. We will now consider progressive application of these constraints.

If each queen is constrained to produce exactly one daughter, but contributions along the fm and mf pathways are random, then the effective population size is derived in case 2 of Table 1 as

$$Ne_{(2)} = \frac{9MF}{3M + 2F}. \tag{2.4}$$

The efficiency of this scheme relative to case 1 for $M = xF$ (where x, the sex ratio, is not necessarily integral) is

$$Ne_{(2)}/Ne_{(1)} = \frac{4M + 2F}{3M + 2F} = \frac{4x + 2}{3x + 2} \tag{2.5}$$

which approaches 4/3 for large x.

If each queen contributes one daughter queen and exactly n drones to the next generation, then variance of family size along both the ff and fm pathways is zero. If semen is collected from all $M(=nF)$ drones and thoroughly mixed (e.g. by the technique of Kaftanoglu and Peng 1980a, b) before samples are used to artificially inseminate each of the virgin daughter queens, then each male will have an equal chance of contributing along the mf pathway. The variance of family size along this pathway will be the Poisson value of F/M. The effective population size for this set of constraints is derived in case 10, Table 1, and

$$Ne_{(10)} = \frac{9MF}{3M + F}. \tag{2.6}$$

For $M = nF$, the efficiency of this scheme relative to case 1 is

$$Ne_{(10)}/Ne_{(1)} = \frac{4n + 2}{3n + 1} \tag{2.7}$$

which approach 4/3 for large n.

The current practice in artificial insemination programmes for bees (Laidlaw 1979) is to collect semen from a sample of drones just before insemination of each virgin queen, rather than to prepare it in bulk as described above. This technique allows for an additional partial constraint to be applied to family size along the mf pathway. If the sample of drones is chosen randomly from the $M = nF$ available, one only of the random drones will produce a single daughter (if the ff pathway is constrained) and the rest will produce no daughters, as opposed to the previous situation where a drone could conceivably produce anything from 0 to F daughters. Now, the Bernoulli variance of family size is appropriate for the mf pathway and the effective population size is derived in case 14

$$Ne_{(14)} = \frac{9MF}{3M - F}. \tag{2.8}$$

The efficiency of this design relative to case 1 is

$$Ne_{(14)}/Ne_{(1)} = \frac{4n + 2}{3n - 1}. \tag{2.9}$$

For small values of sex ratio, this system is more efficient than case 2 and case 10, but for increasing n, the three methods approach the same efficiency. For a sex ratio likely to apply in practice $n = x = 8$, the efficiencies of case 2 and case 10 relative to case 14 are 88.5% and 92%. Clearly if constraining the fm and mf pathways involves additional costs or management difficulties, it is unlikely to be justified in practice.

Page and Laidlaw (1982) obtain their formula 5 for a situation of queen-daughter replacement and random mating, with either random selection of drones for artificial insemination or open mating, presumably in an isolated mating yard. This procedure will lead to a Bernoulli variance of family size along the mf pathway, since the drones are restricted to producing one daughter at most or none at all. The relevant effective population size is presented in case 12 Table 1. If semen for artificial insemination were pooled and mixed from all drones to be used in matings, then $\sigma^2 mf = F/M$ (Poisson), as in case 2 Table 1 and then

$$Ne = \frac{9MF}{3M + 2F}$$

as in their formula 5.

Discussion

The application of constraints on family size along the three pathways of inheritance of sex linked genes leads to substantial increases in effective population size. The most efficient design for increasing autosomal effective population size is that in which each male parent produces a son and a daughter and each female parent likewise contributes a son and a daughter to the next generation. This design doubles autosomal effective population size. Concomittantly, this design will lead to a tripling of sex linked effective population size (formulae 2.1 and 2.2).

The equivalent design for haplo-diploids, in which each male parent produces one daughter, and each female parent contributes one son and one daughter, cannot be used in honeybees since single drone inseminations are not practicable. Singly inseminated queens are not capable of supporting full strength colonies.

Several designs for control populations of honeybees are possible, which for large male-to-female sex ratios tend to increase effective population size by about 1/3 (cases 2, 10 and 14 in order of increasing efficiency) (Page and Laidlaw 1982). The more efficient designs may involve greater cost or management difficulties for little increase in Ne, and choice among them would have to be made on criteria other than efficiency. For example, the bulk mixing of semen in cases 2 and 10 would eliminate the potential problems of incomplete mixing of semen within the spermatheca and non-random expulsion of excess semen (Oldroyd and Moran, in preparation) which might occur with case 14, where semen doses are not homogenised.

Selection programmes which severely restrict the number of breeding queens maintained in a closed line of honeybees are not feasible because of problems caused by the sex allele system (Page and Marks 1982). Briefly, inbreeding increases the probability of homozygosity for balanced lethal sex alleles and eventually leads to such a decrease in colony viability that extinction of colonies, and eventually the line, becomes likely. Page and Marks (1982) have established by computer simulation the constraints on number of queens to be maintained in closed lines when male and female replacements are chosen randomly. They have not considered the effects of selection nor of constraints on family size.

Within-family selection is not a particularly efficient form of selection, but it does have many of the minimum drift advantages of control populations and may be the most appropriate type of selection for honeybees, given the sex allele problem. It is of interest, therefore to ask how the effective population size formulae presented in this paper apply to within-family selection programme in bees. A design for such a programme as follows:

1 ff pathway. Each of the F queens produces q daughters, all of which are inseminated with semen from a homogeneous pool. Colony performance is evaluated and the best 1 of the q daughters is selected. Each queen contributes exactly one daughter to the next generation and therefore $\sigma^2 ff = 0$.

2 fm pathway. Each of the F selected queens contributes exactly $a \times q$ drones (where a is the imsemination factor and equals the equivalent number of single drone semen doses used in each artificial insemination), from which semen is collected and pooled. Since each queen contributes exactly the same number of sons, $\sigma^2 fm = 0$.

3 mf pathway. There is no basis for selection among the M = a.q.F drones. Their contribution to the next generation of queens would be completely random ($\sigma^2 mf = F/M$), except that the within-family selection of replacement daughter queens will cause over or under-representation of daughters of particular drones. Hence the variation of family size along this pathway will be greater than the Poisson value.

Only the case in which all semen is pooled and homogenised is considered. This is because selection among queens will be based on colony performance. The performance of a colony depends equally on contributions by both the queen and the drones to the worker population. Semen mixing would standardise the drone contribution to colony performance, and the differences between colonies should be due only to differences in breeding value between queens as well as random environmental deviations. The only difference between this scheme and control population case 10 is that the variance of family size along the mf pathway is greater than the Poisson value and thus formulae 2.6 for $Ne_{(10)}$ will overestimate effective population size. The magnitude of this overestimation is difficult to gauge. However, for a = 8, q = 3 and F = 50, there will be 1,200 drones represented in the semen pool. The probability of two daughters of the same drone even being present among the 150 unselected daughters is negligibly small, although the probability of choosing daughters of brother drones is more substantial.

Thus for honeybees, within-family selection of this sort will produce an effective population size which approaches the $Ne_{(10)}$ value of $\frac{9MF}{F+3M}$. Of course, if a 1:1 sex ratio were possible in bees, within-family selection could be carried out such that the $Ne_{(16)}$ value of $\frac{9MF}{M+F}$ could be maintained. The effective population size would then be 2.25 times the total number of breeding males and females.

Response obtained to such a within-family selection scheme in honeybees can be predicted from the following formula (Moran and Oldroyd, in preparation). (see Falconer 1981, p 211 for diploid case).

$$R = \frac{2}{3} i \sigma_p h^2 (1 - r) \sqrt{\frac{n-1}{n(1-t)}}$$

where

σ_p = phenotypic standard deviation
h^2 = heritability (narrow sense)
r = relationship between family members
n = number of family members
t = correlation of phenotypic values of members of families, and equals rh^2 where no common environmental factors boost the correlation.

The factor of 2/3 arises because zero selection pressure can be exerted for characteristics of economic importance, such as honey production, along the mf pathway. It is also based on the assumption that equal selection pressure is applied along the fm and ff pathways; a situation which would arise if the selected queens were used both as queen mothers and drone mothers. The relative efficiency of within-family selection to a conventional scheme is determined by the factor $(1 - r)\sqrt{\frac{n-1}{n(1-t)}}$ and will be around 75% or less depending on family size and heritability.

Conventional mass selection programmes lead to a more rapid rate of inbreeding/drift which is likely to lead to a more rapid breakdown of the theoretical, predicted responses to selection. Within-family selection avoids the increased inbreeding/drift due to increased variance in family size, but at the cost of a decrease in the rate of response to selection. Given the particularly adverse consequences of inbreeding in honeybees associated with the sex allele system, within-family selection is a necessary compromise.

Acknowledgements. I would like to thank my colleague, Dr. Frank Nicholas, for advice and constructive criticism during the manuscript preparation, and particularly for directing me to the crucial Pollak (1980) paper. Ben Oldroyd initiated my interest in bee genetics which led to the development of this paper. Several improvements to the final form of the paper are the result of helpful comments from two anonymous referees.

References

Crow JF, Kimura M (1970) An introduction to population genetics theory. Harper and Row, New York

Falconer DS (1981) Introduction to quantitative genetics, 2nd edn. Longman, London New York

Gowe RW, Robertson A, Latter BDH (1959) Environment and poultry breeding problems. 5. The design of poultry control strains. Poult Sci 38:462–471

Hill WG (1972) Estimation of genetic change. 1. General theory and design of control populations. Anim Breed (Abstr) 40:1–15

Kaftanoglu O, Peng YS (1980a) A new syringe for semen storage and instrumental insemination of queen honeybees. J Apic Res 19:73–76

Kaftanoglu O, Peng YS (1980b) A washing technique for collection of honeybee semen. J Apic Res 19:205–211

Laidlaw HH (1979) Contemporary queen rearing. Dadant and Sons, Hamilton, Ill

Latter BDH (1959) Genetic sampling in a random mating population of constant size and sex ratio. Aust J Biol Sci 12:500–505

Moran C, Oldroyd BP (in preparation) Genetic improvement of honeybees in closed populations

Moran PAP, Watterson GA (1959) The genetic effects of family structure in natural populations. Aust J Biol Sci 12:1–15

Oldroyd BP, Moran C (1983) Heritability of worker characters in the honeybee (*Apis mellifera*). Aust J Biol Sci 36:323–332

Ollivier L (1973) Le calcul de l'effectif genetique des populations animales. Ann Genet Sel Anim 5:363–368

Page RE, Laidlaw HH (1982) Closed population honeybee breeding. 2. Comparative methods of stock maintenance and selective breeding. J Apic Res 21:38–44

Page RE, Marks RW (1982) The population genetics of sex determination in honeybees: random mating in closed populations. Heredity 48:263–270

Pollak E (1980) Effective population numbers and mean times to extinction in dioecious populations with overlapping generations. Math Biosci 52:1–25

Wright S (1933) Inbreeding and homozygosis. Proc Natl Acad Sci USA 19:411–420

Wright S (1939) Statistical genetics in relation to evolution. In: Actualites scientifiques et industrielles. Exposes de Biometrie et de la statistique biologique XIII, no. 802. Hermann et lie, Paris, pp 5–64

Wright S (1969) Evolution and the genetics of populations, vol 2. The theory of gene frequencies. University of Chicago Press, Chicago London, pp 204–220