Effective population size in social Hymenoptera with worker-produced males

Robin E. Owen* and
A. R. G. Owen†

* Department of Biological Sciences,
University of Calgary, Alberta, Canada T2N 1N4.
† Department of Zoology University of Toronto,
Toronto, Ontario, Canada M5S 1A1.

In the social Hymenoptera, which have haplodiploid inheritance, a proportion, ψ of the (haploid) males can be produced by the workers. It is shown that, for the special case where each laying worker produces exactly one male that survives to maturity and mates, the variance effective population size $N_{e(v)} = (3 - \psi)^2 FM/(2F + (2 - \psi)^2 M)$, where F and M are, respectively, the number of queens and males in the population. If the sex ratio is unity or female biased then N_e is reduced if there are worker-produced males, however with male biased sex ratios N_e is increased compared to its value with $\psi = 0$. An alternative situation, in which laying workers can each produce more than one male offspring, was investigated using computer simulations. In this case worker-produced males reduce $N_{e(v)}$ regardless of sex ratio, although the effect is relatively the most weak with male biased sex ratios.

A reduction in effective population size due to worker-produced males may contribute to the generally low levels of genetic variation found in the Hymenoptera.

INTRODUCTION

The insect order Hymenoptera is characterized by a haplodiploid genetic system in which, at the phenomenological level, diploid females are derived from fertilized eggs while haploid males arise from unfertilized eggs. In actual fact sex is determined by an underlying genic mechanism (Crozier 1971). Inheritance in haplodiploid species is equivalent to sex-linked inheritance in diploid species.

The Hymenoptera are also characterized by the occurence of eusociality in which there is a reproductive division of labour. Colonies are founded by mated females (queens) who produce daughters (workers and young queens) and males. Workers do not mate, yet in many species retain their ovaries, and under certain conditions can undergo ovarian development and lay unfertilized eggs which develop into males. This production of males by workers is now recognized to be widespread in the higher eusocial Hymenoptera, such as stingless bees, honey bees, bumble bees, vespine wasps and higher ants (for a recent review see Bourke, 1988). The effect that worker-produced males have on the evolution of sociality and optimum sex ratios has received attention (Oster and Wilson, 1978; Aoki and Moody, 1981; Pamilo,

1984; Bourke, 1988). Similarly the effect of worker-produced males on deterministic aspects of hymenopteran population genetics has been investigated (Owen, 1980, 1985a, 1986).

In this paper we examine the influence that worker-produced males have on random genetic drift through effective population size, N_e. There are two contrasting views in the literature on how worker-produced males might affect N_e. Kerr (1975) suggests that given a constant total number of males, then N_e will increase as the proportion of males produced by workers increases. The reasoning followed is that the number of laying workers should be added to the number of queens, to give the total number of "genetically active" females (Contel and Kerr, 1976) in the population, which is then just used in Wright's (1933) standard sex-linked formula for N_e. On the other hand, Crozier (1979) argues that worker-produced males will inevitably reduce effective population size, because an extra round of gametic sampling is introduced each generation and therefore the rate of random genetic drift is increased.

We first examine the simplest case where each laying worker produces exactly one male that survives to maturity and mates. We derive an expression for effective population size and confirm the results using computer simulations. In

this case it turns out that whether worker-produced males increase or decrease N_e depends on the sex-ratio. If the sex-ratio is unity or female biased then N_e is reduced by the presence of worker-produced males, however with male biased sex ratios N_e is increased. Next we consider one alternative scenario in which laying workers produce different numbers of males that survive to maturity and mate. In this particular case we find using computer simulations, that worker-produced males reduce N_e regardless of sex ratio, although the effect is weakest with male biased sex ratios.

ONE OFFSPRING PER LAYING WORKER

The case considered here includes that of ordinary sex-linkage and of haplodiploid inheritance.

Assume that in one generation F queens mate at random with M (haploid) males and their offspring consist of F queens, $(1 - \psi)M$ males and a number of workers. Furthermore it is supposed that ψM of these workers each produce exactly one male by parthenogenesis that survives to maturity and mates. The quantity ψ is the proportion of males produced by the workers which can vary from zero to one (Owen, 1980). Thus the mating population in the next generation still consists of F queens and M males. A population of fixed size is postulated with numbers F and M constant from generation to generation. Consider a single locus with alleles A_1 and A_2 at frequencies in the first generation denoted as follows:

	A_1	A_2
Queens	p_f	q_f
Workers	p_w	q_w
Queen-produced males	p_{m1}	q_{m1}
Worker-produced males	p_{m2}	q_{m2}
All males	p_m	q_m

The actual gene frequency of A_1 among queens as determined by gene count is

$$Fp_f = \text{(number of } A_1A_1 \text{ queens)}$$
$$+ \tfrac{1}{2}\text{(number of } A_1A_2 \text{ queens)},$$

similarly the actual gene frequency of A_1 among queen-produced males given by

$$p_{m1} = \frac{\text{(number of } A_1 \text{ queen-produced males)}}{(1 - \psi)M}.$$

In the next generation a queen or a worker belongs to the genotypes A_1A_1 and A_1A_2 with respective probabilities $p_f p_m$ and $(p_f q_m + q_f p_m)$, thus the expected values of p'_f and p'_w are both equal to

$$Ep'_f = Ep'_w = p_f p_m + \tfrac{1}{2}(p_f q_m + q_f p_m)$$
$$= \tfrac{1}{2}(p_f + p_m). \tag{1}$$

The conditional variance of p'_f is $V(p'_f)$, where

$$FV(p'_f) = p_f p_m + (\tfrac{1}{2})^2 (p_f q_m + q_f p_m) - E(p'_f)^2$$
$$= p_f p_m + \tfrac{1}{4}(p_f + p_m - 2p_f p_m) - \tfrac{1}{4}(p_f + p_m)^2$$
$$= \tfrac{1}{4}(p_f q_f + p_m q_m),$$

i.e.,

$$V(p'_f) = (p_f q_f + p_m q_m)/4F. \tag{2}$$

For queen-produced males $Ep'_{m1} = p_f$, and the variance $V(p'_{m1})$ is just that of the proportion of A_1 gametes in a total of $(1 - \psi)M$ gametes when each gamete has, independently of the others, probabilities p_f of being A_1 and q_f of being A_2. Thus

$$V(p'_{m1}) = p_f q_f / (1 - \psi)M, \tag{3}$$

the binomial variance.

The actual gene frequency of A_1 among worker-produced males equals (number of A_1 worker-produced males)/ψM. This is the proportion of A_1 genes in a sample of ψM genes obtained by randomly drawing one gene from each of the ψM workers who each give rise to a male. Each such gene has independently of the others in the sample, a probability of being A_1 equal to the expected gene frequency of A_1 among the ψM workers. This is therefore $\tfrac{1}{2}(p_f + p_m)$ as in equation (1). Moreover the number of ψM genes in the ψM males produced by workers is binomially distributed. Therefore the variance for p'_{m2} is

$$V(p'_{m2}) = \tfrac{1}{2}(p_f + p_m)[1 - \tfrac{1}{2}(p_f + p_m)]/\psi M$$
$$= [2p_f q_f + 2p_m q_m + (p_f - p_m)^2]/4\psi M. \tag{4}$$

A more formal proof of this formula, whose validity is crucial to the argument, is given in the Appendix.

The expectation is, of course,

$$Ep'_{m2} = \tfrac{1}{2}(p_f + p_m). \tag{5}$$

The gene frequency of A_1 among all males in the next generation is

$$p'_m = (1 - \psi)p'_{m1} + \psi p'_{m2}$$

so that

$$Ep'_m = (1 - \psi)Ep'_{m1} + \psi Ep'_{m2}$$
$$= [(2 - \psi)p_f + \psi p_m]/2. \tag{6}$$

Consequently, if we let

$$\bar{p} = [(2 - \psi)p_f + p_m]/(3 - \psi),$$

then

$$E\bar{p}' =$$

$$[\tfrac{1}{2}(2 - \psi)(p_f + p_m) + \tfrac{1}{2}(2 - \psi)p_f + \tfrac{1}{2}\psi p_m]/(3 - \psi)$$

$$= [(2 - \psi)p_f + p_m]/(3 - \psi) = \bar{p}.$$

Thus $\bar{p}$ is conserved in the sense of expectation; unconditional expectations satisfy

$$E\bar{p}' = E\bar{p} = [(2 - \psi)p_{f0} + p_{m0}]/(3 - \psi) = \bar{p}_0$$

where p_{f0} and p_{m0} are initial (deterministic) values and

$$\bar{p}_0 = [(2 - \psi)p_{f0} + p_{m0}]/(3 - \psi).$$

Also

$$E(p_f' - p_m') = \tfrac{1}{2}(p_f + p_m) - (1 - \tfrac{1}{2}\psi)p_f - \tfrac{1}{2}\psi p_m$$

$$= [-\tfrac{1}{2}(1 - \psi)](p_f - p_m).$$

It follows that for unconditional expectations

$$[E(p_f - p_m)]_t = [-\tfrac{1}{2}(1 - \psi)]^t(p_{f0} - p_{m0}), \quad (7)$$

so that as t approaches infinity p_f and p_m both converge in probability to the same value $\bar{p}_0$. Therefore it can be seen that the deterministic results (Owen, 1980) hold on average, and that these results parallel those of the standard sex-linked case (Avery, 1984).

The variance of p_m' is from equations (3) and (4),

$$V(p_m') = (1 - \psi)^2 V(p_{m1}') + \psi^2 V(p_{m2}')$$

$$= (1 - \psi)^2 \frac{p_f q_f}{(1 - \psi)M}$$

$$+ \frac{\psi^2[2p_f q_f + 2p_m q_m + (p_f - p_m)^2]}{4\psi M}$$

$$= \frac{(2 - \psi)}{2M} p_f q_f + \frac{\psi}{2M} p_m q_m + \frac{\psi}{4M}(p_f - p_m)^2. \quad (8)$$

Effective population size can be obtained from the variance of the change in gene frequency per generation (Avery, 1984; Crow and Denniston, 1988). The variance of the mean gene frequency $\bar{p}'$, conditional on the gene frequencies in the previous generation, $V(\bar{p}'|p_f,p_m)$ is equated to $\bar{p}\bar{q}/(2N_{e(v)})$ (Avery, 1984). Thus the variance effective population size (Crow and Denniston, 1988) is defined as

$$N_{e(v)} = \bar{p}\bar{q}/2V(\bar{p}'|p_f,p_m). \quad (9)$$

The variance of $\bar{p}$ satisfies

$$(3 - \psi)^2 V(\bar{p}'|p_f,p_m) = (2 - \psi)^2 V(p_f') + V(p_m')$$

$$= (2 - \psi)^2 \frac{p_f q_f + p_m q_m}{4F}$$

$$+ \frac{2 - \psi}{2M} p_f q_f$$

$$+ \frac{\psi}{2M} p_m q_m$$

$$+ \frac{\psi}{4M}(p_f - p_m)^2, \quad (10)$$

as given by equations (2) and (8). At a late stage $E[(p_f - p_m)']$ is small, as therefore also $(p_f - p_m)$ must be, so that $p_f \cong p_m \cong \bar{p}$ as the deterministic and stochastic results suggest. Thus substituting in equation (10) gives,

$$(3 - \psi)^2 V(\bar{p}'|p_f,p_m) \cong (2 - \psi)^2 \frac{\bar{p}\bar{q}}{2F}$$

$$+ \frac{2 - \psi}{2M} \bar{p}\bar{q} + \frac{\psi}{2M} \bar{p}\bar{q}$$

$$= \left(\frac{(2 - \psi)^2}{2F} + \frac{1}{M}\right) \bar{p}\bar{q}$$

Therefore

$$2V(\bar{p}'|p_f,p_m)/\bar{p}\bar{q} \cong \left(\frac{(2 - \psi)^2}{2F} + \frac{1}{M}\right)\frac{2}{(3 - \psi)^2}$$

$$= \frac{2F + (2 - \psi)^2 M}{FM(3 - \psi)^2}.$$

Thus, using equation (9), gives

$$N_{e(v)} = \frac{(3 - \psi)^2 FM}{2F + (2 - \psi)^2 M}, \quad (11)$$

which reduces correctly to $9FM/2(F + 2M)$ when $\psi = 0$, this being Wright's (1933) result, derived by the method of path coefficients.

We do not give the (fairly extensive) calculations here, however we have verified that the expression in equation (11) is the correct approximation with large F and M to the value obtained for the effective population size with respect to rate of approach to gene fixation, called "inbreeding effective population size" in Avery (1984) and Crow and Denniston (1988). This is done by forming iterative equations for the three quantities $E(p_f, q_f)$, $E(p_m, q_m)$ and $E[(p_f - p_m)^2]$. Moreover our expression is also the correct approximation to the value for effective population size defined

with respect to mutation. We have obtained exact solutions for each of the above problems. For $\psi = 0$ our exact equation relating to effective population size in respect to gene fixation reduces correctly to that given in Avery (1984); his equation (4·5), page 331. It also satisfies other tests relating to situations where F and M are small. Our exact solution for effective population size defined with respect to mutation applies to arbitrary values of ψ and arbitrary levels of mutation, backward or forward. For $\psi = 0$ and weak mutation our results reduce correctly to the approximate ones given in Nagylaki (1981) and Avery (1984).

Table 1 Examples of effective population size in social Hymenoptera with worker-produced males. Estimates of $N_{e(v)}$ are derived from equation (11) and also from computer simulations (for details see text) and equation (13) ($\hat{N}_{e(v)}$). Three values of ψ (the proportion of worker-produced males) are considered for three sex ratios with the total population number kept constant at 200. The value $\psi = 0$ corresponds to the standard sex-linked case

| | F:M (number) | | | | | |
| | 150:50 | | 100:100 | | 50:150 | |
ψ	$N_{e(v)}$	$\hat{N}_{e(v)}$	$N_{e(v)}$	$\hat{N}_{e(v)}$	$N_{e(v)}$	$\hat{N}_{e(v)}$
0	135·0	130·7	150·0	155·3	96·4	92·2
0·5	113·6	115·9	147·1	149·9	107·1	106·0
1·0	85·7	93·3	133·3	130·6	120·0	116·3

Effective population size given by equation (11) was compared to that estimated from computer simulations of random genetic drift (table 1). The variance in gene frequency over 150 replicate populations was calculated for 20 generations. In each replicate an initial gene pool was constituted with alleles A_1 and A_2 at equal frequencies of $p_f = p_m = 0.5$. Exactly ψM worker genotypes were formed by picking gametes (sperm and eggs) at random, and each gave rise to a single male. A single gamete (A_1 or A_2) was picked at random with equal probability from heterozygous workers to derive their male offspring. Queen-produced males were picked from the gene pool with A_1 at probability p_f and A_2 at probability q_f. Young queen genotypes were formed by picking gametes at random and the new gene frequencies calculated from the genotype frequencies by gene counting. The new gene frequencies in males were calculated by combining the queen—and worker-produced males. The process was then repeated using the new frequencies p_f' and p_m'.

It is well known that at an autosomal locus the variance in gene frequency at generation t,

$(\sigma^2 p)_t \cong p_0 q_0 (1 - e^{-t/2N_e})$ where p_0 and q_0 are the initial gene frequencies. Now since $V(\bar{p}' | p_f, p_m) = \bar{p}\bar{q}/2N_{e(v)}$ (equation 9), the following holds approximately

$$(\sigma^2 \bar{p})_t \cong \bar{p}_0 \bar{q}_0 (1 - e^{-t/2N_{e(v)}}), \qquad (12)$$

where $\bar{p}_0$ and $\bar{q}_0$ are the initial mean allele frequencies. Hence effective population size can be estimated from the computer simulations as

$$\hat{N}_{e(v)} \cong - t/\{2 \ln[(\bar{p}_0 \bar{q}_0 - s^2{}_{pt})/\bar{p}_0 \bar{q}_0]\}, \quad (13)$$

where $s^2{}_{pt}$ is the observed variance in gene frequencies over the replicate populations. In the computer simulations $\bar{p}_0 = \bar{q}_0 = 0.5$ and $t = 20$. Effective population size was determined for three values of ψ for each of three sex ratios (table 1). The estimates of $N_{e(v)}$ derived from the formula (equation(11)) and the computer simulations are in close agreement. The effect of worker-produced males is to reduce N_e if the sex ratio is unity or female biased, however N_e is increased, compared to its value with $\psi = 0$, if the sex ratio is male biased. The proportional changes in N_e for various values of ψ and different sex ratios are given in table 2. It can be seen that the proportional reduction in N_e with female biased ratios is greater than the proportional increase with male biased ratios.

Table 2 Effective population size in social Hymenoptera as a function of sex ratio and the proportion of males produced by workers (ψ). The ratio of $N_{e(v)}$ for any particular value of ψ to that when $\psi = 0$ is given for each sex ratio. $N_{e(v)}$ was calculated using equation (11)

| | Sex Ratio $F:M$ | | | | | | |
ψ	3:1	2:1	1·5:1	1:1	1:1·5	1:2	1:3
0	1·00	1·00	1·00	1·00	1·00	1·00	1·00
0·25	0·93	0·95	0·97	1·00	1·02	1·03	1·05
0·50	0·84	0·89	0·93	0·98	1·03	1·07	1·11
0·75	0·74	0·81	0·86	0·95	1·04	1·10	1·18
1·0	0·63	0·71	0·78	0·89	1·02	1·11	1·24

MORE THAN ONE OFFSPRING PER LAYING WORKER

The case analyzed in the previous section represented one extreme situation in which each laying worker produced exactly one male that survived to maturity and mated. This might, of course, be a realistic representation of what occurs in nature, but at the moment we have no way of knowing whether this is true or not. A more realistic scenario would be if each laying worker produced a

different number of males. We can envisage that there are a number of alternative ways in which this could occur. For instance each worker might produce a random number of males. However we chose to model by computer simulation a simple case in which equal numbers of workers produce either 1, 2, 3, 4, or 5 males. This introduces a deterministic element to the problem, but serves to illustrate very well the effect of unequal family size.

The simulations were done in the same way as described earlier except for the generation of worker-produced males. This time the genotypes of the laying workers were constituted by picking gametes at random as before, but now each worker gave rise to between one and five male offspring. Gametes from heterozygous workers were picked at random with equal probability and if these gave rise to more than one male then this random sampling was repeated the appropriate number of times. Consider as an illustration an example with an equal sex ratio of $100F:100M$ in the population and $\psi = 0.9$; thus 90 of the males would be worker-produced. Therefore in the simulation six workers would each give rise to one male, another six would each give rise to two males, and so on until a final six would each give rise to five males, and the genotypes of all of these 30 laying workers would have been picked at random.

The results of some simulations are given in table 3. The estimated effective population size $\hat{N}_{e(v)}$, calculated using equation (13), is shown together with the variance in female allele

Table 3 Results of computer simulations of effective population size in social Hymenoptera where laying workers can each produce more than one male offspring (see text for details). Total population number was set at 200 and the sex ratio and total number of worker-produced males was varied. Simulations were run using 150 replicate populations. The values of mean heterozygosity ($\bar{H}$), the variance of gene frequency in females (s_p^2), and $\hat{N}_{e(v)}$ (from equation (13)) at generation 20 are given in the table

Number $F:M$	Number of worker-produced males	ψ	$\bar{H}$	s_p^2	$\hat{N}_{e(v)}$
100:100	0	0	0.4730	0.0158	153.5
	45	0.45	0.4634	0.0191	125.9
	90	0.90	0.4354	0.0326	71.6
125:75	0	0	0.4708	0.0169	143.2
	75	0.5	0.4591	0.0213	112.2
	150	1.0	0.4174	0.0417	54.8
50:150	0	0	0.4574	0.0237	100.5
	30	0.4	0.4534	0.0254	93.2
	75	1.0	0.4444	0.0295	79.6

frequency and mean heterozygosity (where $H = p_f q_m + q_f p_m$) after 20 generations. Worker-produced males in this case reduce effective population size regardless of sex ratio. However the effect is least with male biased sex ratios. Note that when the sex ratio is unity or female biased the proportional reduction in N_e is considerably greater than in the case of one male per laying worker. For example with a 1:1 sex ratio and $\psi = 1.0$, N_e is reduced by 11 per cent when each worker produces a single male (tables 1 and 2), however when some workers produce more than one male the reduction in N_e is 29 per cent (table 3).

It is of course possible to theoretically extend the formula of equation (11) to cover all cases of variable size of families produced by queens and workers, and we have indeed done this. The algebraic foundations are however somewhat lengthy and as the main thrust of this paper is in the direction of fact finding rather than mathematical exposition, in the interests of brevity we have given preference to illustrating the problem via simulation.

DISCUSSION

In this paper we have provided the first quantitative treatment of the problem of effective population size in the social Hymenoptera that have worker-produced males. We derived an explicit formula for N_e for the special case where each laying worker produces exactly one male that survives to maturity and mates. This formula showed that worker-produced males reduce N_e when population sex ratios are either unity or female biased, but increase N_e when sex ratios are male biased. Thus a reduction in N_e is not an inevitable consequence of introducing an extra round of gametic sampling each generation. Computer simulations of a perhaps more realistic situation, where some laying workers produce more than one male offspring, showed a reduction in N_e regardless of sex-ratio. Therefore it seems reasonable to conclude that worker-produced males will generally tend to reduce effective population size rather than increase it, as argued by Crozier (1979).

We have assumed that the proportion of worker-produced males (ψ) is constant from one generation to the next. In fact it may fluctuate stochastically from generation to generation, for example in species where workers produce males in orphaned colonies and the frequency of orphaning (queen loss) depends on ecological conditions.

It is well known that when a population is subject to variation in numbers then the effective population size is approximated by the harmonic mean of the successive generation sizes (Wright, 1938). Consider now the situation where population number and sex ratio is fixed but ψ varies each generation. Then applying the same principle we see that approximately,

$$1/\bar{N}_e = (1/t)\sum(1/N_{ei})$$

where $\bar{N}_e$ is the harmonic mean of the effective population sizes (N_{eis}) dependent on ψ. Since the harmonic mean tends to be dominated by the smallest terms, it follows that N_e will be further reduced by fluctuating values of ψ relative to its value if there were no worker-produced males.

The change in effective population size caused by having worker-produced males can be quite substantial (table 1), although the effect tends to be more pronounced when some workers can each produce more than one male offspring. Therefore, as far as 1:1 and female biased sex ratios are concerned, equation (11) can be thought of as giving a minimum estimate of the reduction in $N_{e(v)}$ that will occur for any particular value of ψ recorded.

The significant effect that worker-produced males can have on effective population size has a bearing on the maintenance of genetic variability in the Hymenoptera. In general Hymenoptera have significantly lower levels of genetic variation, at least at electrophoretically detectable loci, than comparable diploid insects (Berkelhammer, 1983; Graur, 1985; Owen 1985b). Gene loci in haplodiploids are predicted to have lower mean heterozygosities than loci in diploids whether selection or mutation-drift maintain variability (Avery, 1984). Variation at loci possessing selectively neutral alleles depends on random drift and mutation. N_e is lower at haplodiploid loci than comparable diploid loci (unless $F > 7M$, Crozier, 1976) and so the effect of drift is greater. Worker-produced males may, therefore, generally enhance this effect. Also N_e enters into many formulae incorporating drift, mutation and selection (see Avery, 1984) including probability of and time to fixation, which thus can be affected indirectly by worker-produced males.

REFERENCES

AOKI, K. AND MOODY, M. 1981. One- and two-locus models of the origin of worker behaviour in Hymenoptera. *J. Theor. Biol.*, 89, 449–474.

AVERY, P. J. 1984. The population genetics of haplo-diploids and x-linked genes. *Genet. Res.*, 44, 321–341.

BERKELHAMMER, R. C. 1983. Intraspecific genetic variation and haplodiploidy, eusociality and polygyny in the Hymenoptera. *Evolution*, 37, 540–545.

BOURKE, A. F. G. 1988. Worker reproduction in the higher eusocial Hymenoptera. *Q. Rev. Biol.*, 63, 291–311.

CONTEL, E. P. B. AND KERR, W. E. 1976. Origin of males in *Melipona subnitida* estimated from data of an isozymic polymorphic system. *Genetica*, 46, 271–277.

CROW, J. F. AND DENNISTON, C. 1988. Inbreeding and variance effective population numbers. *Evolution*, 42, 482–495.

CROZIER, R. H. 1971. Heterozygosity and sex determination in haplo-diploidy. *Amer. Nat.*, 105, 399–412.

CROZIER, R. H. 1976. Counter-intuitive property of effective population size. *Nature*, 262, 384.

CROZIER, R. H. 1979. Genetics of sociality. In Herman, H. R. (ed.) *Social Insects, Volume I*, Academic Press, New York, pp. 223–286.

GRAUR, D. 1985. Gene diversity in Hymenoptera. *Evolution*, 39, 190–199.

KERR, W. E. 1975. Evolution of the population structure in bees. *Genetics*, 79, 73–84.

NAGYLAKI, T. 1981. The inbreeding effective population number and the expected homozygosity for an X-linked locus. *Genetics*, 97, 731–737.

OSTER, G. AND WILSON, E. O. 1978. *Caste and Ecology in the Social Insects*. Princeton University Press, Princeton, N. J., U.S.A.

OWEN, R. E. 1980. Population genetics of social Hymenoptera with worker-produced males. *Heredity*, 45, 31–46.

OWEN, R. E. 1985a. The opportunity for polymorphism and genic variation in social Hymenoptera with worker-produced males. *Heredity*, 54, 25–36.

OWEN, R. E. 1985b. Difficulties with the interpretation of patterns of genetic variation in the eusocial Hymenoptera. *Evolution*, 39, 201–205.

OWEN, R. E. 1986. Colony-level selection in the social insects: single locus additive and nonadditive models. *Theor. Pop. Biol.*, 29, 198–234.

PAMILO, P. 1984. Genetic relatedness and evolution of insect sociality. *Behav. Ecol. Sociobiol.*, 15, 241–248.

WRIGHT, S. 1933. Inbreeding and homozygosis. *Proc. Natl Acad. Sci., USA*, 19, 411–420.

WRIGHT, S. 1938. Size of population and breeding structure in relation to evolution. *Science*, 87, 430–431.

APPENDIX

Let m workers only each produce one male offspring, and m_1, m_2, m_3 be the respective numbers of such workers of genotypes A_1A_1, A_1A_2, A_2A_2, where $m_1 + m_2 + m_3 = m$. Conditional on p_f and p_m, the respective probabilities of a worker being of one of these genotypes are $\theta_1, \theta_2, \theta_3$, Thus the probability of the observed distribution is the multinomial term $m! \; \theta_1^{m_1} \theta_2^{m_2} \theta_3^{m_3} / m_1! m_2! m_3!$. Conditional on this partition m_1, m_2, m_3 of m the proba-

bility that there will be $(m_1 + m_1^*)$ males of genotype A_1 and $(m_3^* + m_3)$ males of genotype A_2 is just the binomial probability $m_2!/2^{m_2}m_1^*!m_3^*!$, with $m_1^* + m_3^* = m_2$. Thus, conditional on p_f and p_m the probability of $m_1 + m_1^*$ males A_1 and $m_3^* + m_3$ males A_2 is just

$$\frac{m!}{m_1!\,m_2!\,m_3!}\frac{m_2!}{m_1^*!\,m_3^*!}\,\theta_1^{m_1}\left(\tfrac{1}{2}\theta_2\right)^{m_1^*}\left(\tfrac{1}{2}\theta_2\right)^{m_3^*}\theta_3^{m_3}$$

which is the multinomial probability for four classes totalling m to be distributed as m_1, m_1^*, m_3^*, m_3 when the class probabilities are $\theta_1, \tfrac{1}{2}\theta_2$, $\tfrac{1}{2}\theta_2$, θ_3. We require the probability for the m objects (*i.e.*, genotypes) of type A_1 or A_2 to be distributed as s_1, s_2 where $s_1 = m_1 + m_1^*$, $s_2 = m_3^* + m_3$. As the result of pooling classes in a multinominal distribution is a multinominal with an appropriately reduced number of classes, it follows that $s_1 = m_1 + m_1^*$ is distributed as a binomial variate with index m and parameters $\theta_1 + \tfrac{1}{2}\theta_2 = \tfrac{1}{2}(p_f + p_m)$, which was to be proved.

If only variances are required without knowledge of the probability distributions, they can be calculated as follows. Thus for parthenogenetically produced haploid males (whether from queens or workers) we can set up the table in which g_1 is the number of A_1 genes in a single male offspring;

Parental genotype	Frequency	Male genotype probabilities	
		A_1 $g_1 = 1$	A_2 $g_1 = 0$
A_1A_1	θ_1	1	0
A_1A_2	θ_2	$\tfrac{1}{2}$	$\tfrac{1}{2}$
A_2A_2	θ_3	0	1

Conditional expectations and variances

	Eg_1	$E(g_1^2)$	$(Eg_1)^2$	$V(g_1)$
A_1A_1	1	1	1	0
A_1A_2	$\tfrac{1}{2}$	$\tfrac{1}{2}$	$\tfrac{1}{4}$	$\tfrac{1}{4}$
A_2A_2	0	0	0	0

Thus $E(g) = E(g^2) = \theta_1 + \tfrac{1}{2}\theta_2$, $(Eg)^2 = (\theta_1 + \tfrac{1}{2}\theta_2)^2$ so that $V(g) = \theta_1 + \tfrac{1}{2}\theta_2 - (\theta_1 - \tfrac{1}{2}\theta_2)^2 = p_f q_f$ for queen-produced males $= \tfrac{1}{2}(p_f + p_m)[1 - \tfrac{1}{2}(p_f + p_m)]$ for worker-produced males, because $\theta_1 + \tfrac{1}{2}\theta_2 = p_f p_m + \tfrac{1}{2}(p_f q_m + q_f p_m) = \tfrac{1}{2}(p_f + p_m)$. The required variances follow on division by the appropriate divisors, $(1 - \psi)M$ for queen-produced males and ψM for worker-produced males.