

# The infinite island model with sex-differentiated gene flow

LARS M. BERG\*†, MARTIN LASCOUX‡§ & PEKKA PAMILO†

†Department of Genetics, Box 7003, Uppsala University, S-750 07 Uppsala and ‡Department of Forest Genetics, Box 7027, Swedish University of Agricultural Sciences, S-750 07 Uppsala, Sweden

Identity measures are derived for the infinite island model with separate sexes and sex-differentiated contribution to gene flow. The concept of effective migration rate,  $m_e$ , is introduced, which describes the genetically effective flow when sexes migrate at different rates, when females migrate as mated or when sexes differ in ploidy, as in the case of X-linked genes or haplodiploidy. The notion of  $m_e$  allows the different cases to be described by the common equation  $F = 1/(1 + 4N_e m_e)$ . When the reproductive sex ratio differs from unity, migration of the rarer sex entails more effective gene flow than migration of the common sex. With female dispersal after mating, as well as with ploidy differences, the effective migration rate departs from the census rate, and migration of diploids and/or mated females normally homogenizes the population more than the migration of haploids and unmated females, as expected. This difference between the effective gene flow and the number of migrants contributes to a possible discordance between direct and indirect estimates of dispersal. At the same time, the expected difference in genetic differentiation between autosomal and X-linked loci opens the possibility for revelation of sex-differentiated migration, which may help resolve such instances.

**Keywords:** genetic differentiation, haplodiploidy, identity measures, island model, sex-biased migration.

## Introduction

The equilibrium between drift and migration is expected to result in a specific level of genetic differentiation (e.g. quantified as  $F_{ST}$ ). As a consequence, studies of genetic differentiation of populations allow indirect estimation of migration (Felsenstein, 1982; Slatkin, 1985), and the amount of differentiation can be predicted under specific migration models if the migration rates are known or under management control. For instance, under the island model (Wright, 1931)

$$F_{ST} = \frac{1}{1 + 4N_e m}, \quad (1)$$

where  $N_e$  is the effective population size and  $m$  is the migration rate. Our aim is to specify this relationship when males and females contribute unequally to gene flow. This can happen when the

migration rates are sex dependent, the timing of migration in relation to mating differs between the sexes and when the sexes are of different ploidy, as in the case of sex chromosomal genes and in species with male-haploid sex determination (e.g. hymenopteran and thysanopteran insects).

Many animal species show a consistent and distinct pattern of sex-dependent migration, such that one sex predominates among migrating individuals. In mammals, for example, migration is often male biased, whereas in birds the reverse is found (Greenwood, 1980) and, among insects, extremely pronounced bias is not uncommon. Similarly, in plants, strong differences in pollen and seed dispersal have been documented (Ennos, 1994). Recently, Wang (1997a,b) considered sex-dependent migration in a diploid system with a finite number of subpopulations, arbitrary distribution of family sizes and partial selfing. The results support earlier work by Prout (1981) and Ennos (1994) in the restricted case of an infinite number of subpopulations, Poisson distribution of family sizes and random mating. In this case, eqn (1) becomes

§Current address: Department of Genetics, Uppsala University, Sweden.

\*Correspondence. E-mail: lars.berg@genetik.uu.se

$$F_{ST} = \frac{1}{1 + 2N_e(m_f + m_m)}, \quad (2)$$

when females disperse before mating. Here,  $m_f$  and  $m_m$  are chosen to denote the migration rates of females and males, respectively. When females disperse after mating

$$F_{ST} = \frac{1}{1 + 2N_e(2m_f + m_m)}. \quad (3)$$

Gene flow in plants often belongs to the latter category, which was consequently termed pollen–seed migration by Prout (1981). If the cases in which dispersing males have new mating opportunities are neglected, the dispersal of males after mating has no effect on the gene pools.

Prout (1981) further suggested that his pollen–seed model (eqn 3) is applicable to male-haploid genetic systems (using wasps as an example), as the haploid males could be seen as equivalent to pollen grains. This analogy has been used when discussing kinship relationships in societies of social and male-haploid bees (Laidlaw & Page, 1988). A honeybee queen can be seen as a hermaphrodite individual and the haploid drones as her enlarged sperm packages. The male-haploid system does not, however, correspond precisely to the dispersal of pollen and seed in a diploid monoecious system. We should allow unequal sex ratios, and we must note that, if the haploid males are considered as pollen, they do not fertilize egg cells of the same, but of the following, generation.

We will here derive the equilibrium values of  $F_{ST}$  in an infinite island model with separate sexes, allowing for sexual differences in ploidy level. We restrict ourselves to the simplified case of random mating and Poisson-distributed fecundities. For the sake of generality and demonstration, we will first examine the diploid cases already considered by others and then extend the results to a case in which either sex is haploid.

## Models

### General outlines

We will adopt largely the same assumptions as in Wright's infinite island model (Wright, 1931): each deme or subpopulation is made up of a constant number of females ( $N_f$ ) and males ( $N_m$ ), there is no difference in offspring sex ratios between migrating and non-migrating individuals, generations are discrete, mating is random and the number of offspring is Poisson distributed among parents of

each sex. We will also assume that mutation rates are low compared with the migration rates and, therefore, can be neglected. We treat the two different migration schemes of female dispersal before mating (henceforth referred to as DBM models) and female dispersal after mating (referred to as DAM).

The coefficient of kinship,  $f$ , of individuals is calculated as the probability of sampling two genes identical by descent from two zygotes produced in a single deme. With random mating  $f_t = F_{t+1}$ , i.e. the inbreeding coefficient of the zygotes of the next generation. Denote by  $P_f$  and  $P_m$  the probability of sampling a zygotic gene originating from a female and a male, respectively. Assuming that the identity by descent of genes is independent of the sex, the value of  $f$  in generation  $t + 1$  becomes

$$\begin{aligned} f_{t+1} = & P_f^2(1 - m_f)^2 \left[ \frac{\pi_f}{N_f} + \left(1 - \frac{1}{N_f}\right) f_t \right] \\ & + P_m^2(1 - m_m)^2 \left[ \frac{\pi_m}{N_m} + \left(1 - \frac{1}{N_m}\right) f_t \right] \\ & + 2P_f P_m f_t (1 - m_f)(1 - m_m), \end{aligned} \quad (4)$$

where  $\pi_f$  ( $\pi_m$ ) denotes the probability of identity of two genes from the same female (male). At the drift–migration equilibrium  $f_{t+1} = f_t = F_{ST}$ , which measures the level of kinship within and differentiation among the demes.

### Diploid models

In the DBM model, both sexes migrate before mating with the migration rates  $m_f$  and  $m_m$ . As half of the genes originate from females and half from males,  $P_f = P_m = 1/2$ , independently of the sex ratio. In the diploid case,  $\pi_f = \pi_m = 1/2 + f_{t-1}/2$ . Noting that

$$N_e = \frac{4N_f N_m}{N_f + N_m},$$

we get from eqn (4), at equilibrium:

$$F_{ST} \approx \frac{1}{1 + 4N_e \frac{m_f + m_m}{2}}. \quad (5)$$

when the second order of migration rates can be neglected. This is the same result as eqn (2). The concept of effective migration is introduced,  $m_e = (m_f + m_m)/2$ , which allows the writing of eqn (5) in the same form as Wright's classic formula

$$F_{ST} = \frac{1}{1 + 4N_e m_e} \quad (6)$$

Equation (5) is also obtained by deriving the first and second moments of allele frequencies under the joint effects of drift and migration (Crow & Kimura, 1970, pp. 439–440) and separating these effects by sex (Appendix).

The model with dispersal after mating (DAM) can easily be derived from the above results. The dispersal of females after mating means that they carry along the sperm received from their mates. The genetic effect of this is that the immigration rate of male genes equals  $m_m + m_f$ , where the male migration takes place before mating and the female migration after it, whereas the migration of female genes remains unaltered. This is true because, when the sex ratio of zygotes carried by migrating females equals that of residents, the male and female zygote migration rates are both equal to the total rate of migrating zygotes, which, provided that the average female only replaces herself, equals the migration rate of mated females. It should be noted that, when females on average give birth to more than one male, DAM induces kin-structured migration of male genes. However, this has no effect on the within-deme identity as long as migration rates are of the order assumed here. Thus, the effective immigration rate can be written as:

$$m_e = \frac{2m_f + m_m}{2} \quad (7)$$

and eqn (6) applied for  $F_{ST}$ . This agrees with eqn (3).

#### Haplodiploid models

In the case of sex chromosomes or haplodiploidy, the contributions of the sexes to the gene pool are unequal. Assuming that the males form the heterogametic/haploid sex, the contributions of males and females in the asymptotic gene pool (Oster *et al.*, 1977) are  $1/3$  and  $2/3$ , respectively (Pamilo, 1991). These values are independent of the population sex ratio and can be used for calculating the identity of genes in the offspring generation in the same way as in the diploid case. Thus, setting  $P_f = 2/3$  and  $P_m = 1/3$ , and noting that, when males are haploid,  $\pi_m = 1$  and

$$N_e = \frac{9N_f N_m}{2N_f + 4N_m}$$

(Wright, 1969), we arrive at:

$$F_{ST} \approx \frac{1}{1 + 4N_e m_e}, \quad (8)$$

where

$$m_e = \frac{2m_f + m_m}{3}$$

under the same assumptions as for eqn (5). If the female is the heterogametic (or haploid) sex, the subscripts *f* and *m* have to be switched.

When the females disperse after mating, it is necessary to distinguish clearly which sex is haploid. Using the same arguments as above, the effective migration rates in the DAM model are:

$$m_e = \frac{3m_f + m_m}{3} \quad (10)$$

when the males are haploid and

$$m_e = \frac{3m_f + 2m_m}{3} \quad (11)$$

when the females are haploid.

#### Evaluating the models

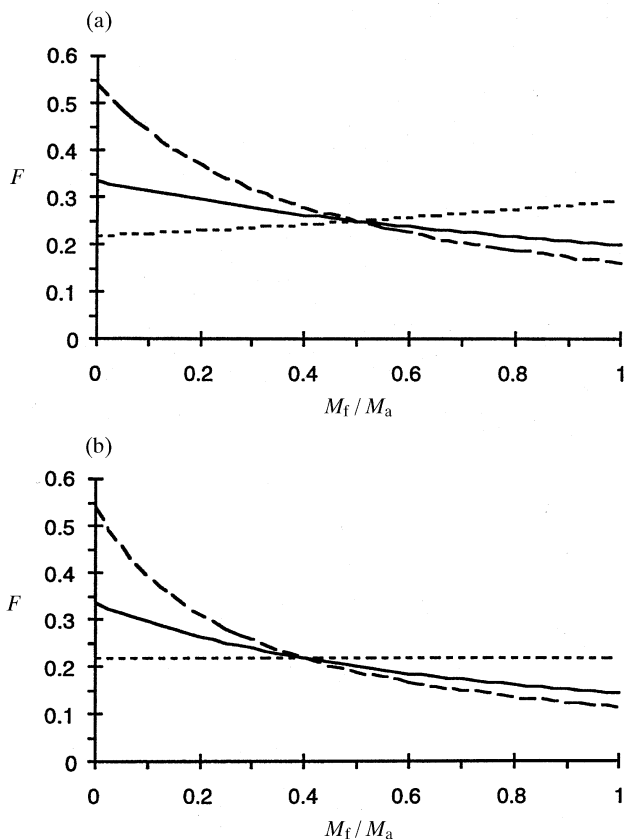
The differentiation between populations is given by eqn (6), and this formula allows examination of the effects of sex ratios (via the effective population sizes) and sex-dependent migration rates. We present the cases in which the actual number of migrants  $M_a = N_f m_f + N_m m_m = 1$ . Figure 1 gives the values of  $F_{ST}$  for haplodiploid populations in the following cases: (i) fixed sex ratios and variable ratio of migrants; and (ii) fixed sex-biased migration and variable reproductive sex ratios for constant total deme sizes. It should be noted that our derivations assume small immigration rates. The numerical presentations can, in some cases, violate this assumption when the sex ratios become extremely biased and the number of migrants of the rarer sex is kept high.

For the diploid model with even sex ratios and dispersal before mating,  $F_{ST} = 0.2$  when there is one migrant per generation ( $M_a = 1$ ), independently of the migrant sex ratio. The other results can now be compared with this basic case (Fig. 1). The haplodiploid models with even sex ratios have  $F_{ST} = 0.2$  when there is one diploid migrant per generation, and  $F_{ST} > 0.2$  when part of the immigration is by haploid individuals. Dispersal after mating always induces

more gene flow and reduces the values of  $F_{ST}$ . The DAM models result in  $F_{ST} < 0.2$  when the females are haploid, or when more than half of the migrants are females when the males are haploid.

When the sex ratios depart from 1:1, migration of the rarer sex homogenizes the populations, and the  $F_{ST}$ -values are smaller than with even sex ratios, whereas the opposite is true when the migrants belong to the commoner sex. The values of  $F_{ST} > 0.2$  when the sex ratio within immigrants equals that within populations and departs from the 1:1 ratio.

The value of  $F_{ST}$  depends directly on the effective gene flow. We define this quantity as  $M_e = N_e m_e$  and compare this value with the actual number of immigrants  $M_a = N_f m_f + N_m m_m$  (Fig. 2). The patterns are, of course, identical to those for  $F_{ST}$ . With even sex ratios,  $M_a \leq M_e \leq 2M_a$  when the females disperse after mating, and  $M_e = M_a/2$  in the male-haploid model when only males disperse. The ratio of  $M_e/M_a$  varies more widely when the sex ratios are uneven.

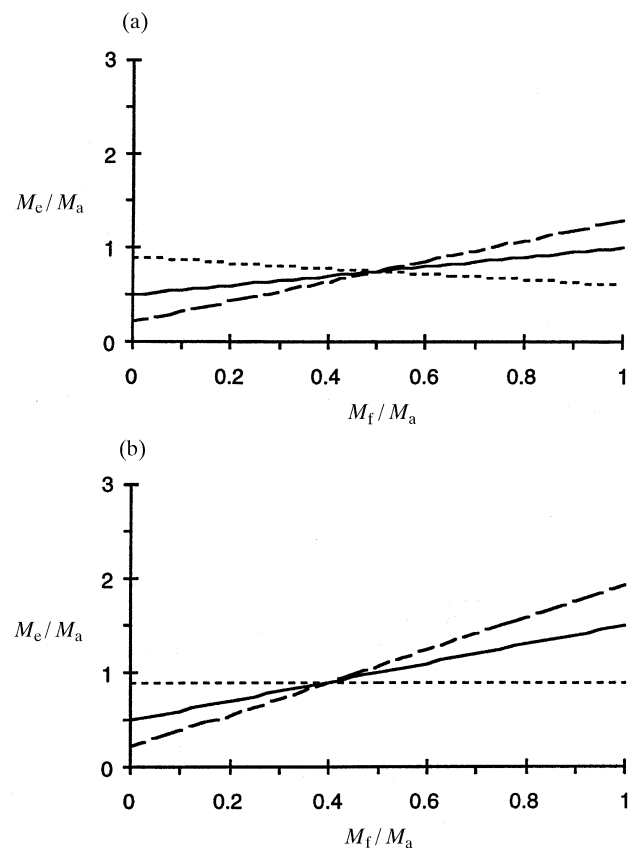


**Fig. 1** Equilibrium  $F$  as a function of the proportion of females among migrants ( $M_f/M_a$ ), for three reproductive sex ratios (females:males): — 1:1, --- 3:1, ..... 1:3, in (a) the male-haploid DBM and (b) the male-haploid DAM model.  $M_a = 1$ .

The possible range depends on how uneven the sex ratios are.

## Discussion

The island model of migration (Wright, 1931) was initially derived for diploid monoecious organisms, and sex-specific migration has been considered for several situations (Prout, 1981; Maruyama & Tachida, 1992; Ennos, 1994; Wang, 1997a,b). As far as we know, ours is the first attempt to apply the island model to sex-chromosomal genes or haplodiploid species. It is pleasing to note that the drift-migration balance for models with unequal sex ratios, sex-specific migration rates and varying ploidy levels can be presented in the same general form (eqn 6) as that of the classic island model (eqn 1). However, the final effect of male or female migration biases on the amount of differentiation depends on the reproductive sex ratios, and the interpreta-



**Fig. 2** Ratio of the number of effective migrants and the census number of migrants ( $M_e/M_a$ ), as a function of the proportion of females among migrants ( $M_f/M_a$ ), for three reproductive sex ratios (females:males): — 1:1, --- 3:1, ..... 1:3, in (a) the male-haploid DBM and (b) the male-haploid DAM model.  $M_a = 1$ .

tion of the product  $N_e m_e$  in terms of the number of migrants is less straightforward.

When the sex ratio is held unbiased, selfing is disallowed, and mutation is ignored, our results for diploid systems agree with the hermaphrodite/unbiased models of Maruyama & Tachida (1992), Ennos (1994) and Prout (1981), but they depart from these when either sex is haploid or when there is a reproductive sexual bias. This is in agreement with Wang's (1997b) findings for the DBM model, and our results for the dioecious DAM model extend his (Wang, 1997a) work in also accounting for varying sex ratios in the case of pollen–seed migration with random mating. Prout (1981) also showed that sex-differentiated migration induces a particular form of outcrossing, which results in heterozygote excesses within local populations and, thus, in negative  $F_{IS}$ -values. This effect is also expected in the case of sexual ploidy differences. Note, however, that the heterozygote excess is negligible when migration rates are at the level assumed in eqn (6).

Our results show, as expected, that genetic differentiation for a given number of migrants is highest when the migrants are of the haploid sex and/or when they belong to the common sex. In general, genetic differentiation of populations is expected to be higher for sex-chromosomal than for autosomal loci, unless the sex ratio and the migration rate are very biased and the migrants belong to the homogametic common sex. The trend is similar for the two models, DBM and DAM, but female dispersal after mating reduces population differences effectively. This effect is particularly evident when the population sex ratios are male-biased.

Prout (1981) suggested that his pollen–seed model, which corresponds to our DAM model for diploid systems, could apply to male-haploid wasps. However, this involves viewing the females as hermaphrodites and disregarding the males in the population number. It also implicitly assumes a sex ratio of unity for the haplodiploid population in question. Because of these requirements, direct comparison of the dynamics of the different equations would be misleading. Instead, we are restricted to the special case of unbiased reproductive ratio, for which the two models agree well. Thus, the 'postponement' of male reproduction appears to have no effect on the equilibrium differentiation of male-haploid populations, and the diploid DAM model can be applied with no problem as long as the reproductive sex ratio is close to unity.

The amount of genetic differentiation under the drift–migration balance reflects the effective gene flow.

The difference between the effective and the absolute number of migrants, demonstrated in this paper, contributes to a possible discordance between estimates of migration obtained by direct methods (by counting dispersing individuals) and by indirect methods (by inference from the genetic structure) (Slatkin, 1985). For example, consider a diploid case in which both the reproductive and migrant sex ratios are skewed in favour of females. Because the effect of female migrants as gene carriers is smaller than that of the males, a simple count of immigrants yields an overestimate of the effective gene flow. In this instance, information about sex ratios and sex-differentiated migration would help to interpret the genetic data in terms of actual numbers. A number of methods have been developed that use a comparison of the differentiation for nuclear and cytoplasmic genes to reveal sexual migration patterns (see, for example, Avise, 1994). The expected difference in genetic differentiation for autosomal and X-linked genes, derived here, opens the possibility for yet another way of estimating sex-biased migration. Such a method would be most useful when the heterogametic (or haploid) sex predominates among migrants, as we then expect the contrast between these types of genes to be most pronounced.

Sex-biased migration and skewed secondary or reproductive sex ratios are common phenomena among animals and plants (e.g. Greenwood, 1980; Ennos, 1994). The two phenomena often occur together. Such cases are of special interest in relation to the present work, because then we normally expect the effect on equilibrium genetic differentiation to be at its greatest. For example, Greenwood (1980, 1983) noted a correlation among birds and mammals between polygyny and male dispersal bias. Thus, we may often expect a female-biased reproductive ratio to co-occur with a male-biased migrant ratio, which means that, using Wright's equation, we would underpredict the equilibrium differentiation from knowledge of migration rates or overestimate the number of migrants from genetic data. Note, however, that if the skewed reproductive ratio is caused by differential mating success, the assumption of Poisson-distributed fecundity is most probably violated. Furthermore, if there are large sexual differences in survival rates, immigrants of the rarer sex may have difficulties succeeding with mating. This could often balance a sexual bias in dispersal, decreasing its genetic effects.

Genetic studies of ants have shown significant differentiation among nearby populations in species that build large colonial networks and apparently

produce new nests largely by budding. Estimates of  $F_{ST}$  are about 0.2 in the species *Myrmica rubra* (Seppä & Pamilo, 1995), *Formica truncorum* (Sundström, 1993) and *F. aquilonia* (P. Pamilo, unpubl. results). If all migration is by males and the sex ratio is  $x$  males per female,  $F_{ST} = 0.2$  corresponds to the number of immigrating males  $N_m m_m = (2 + 4x)/3$  per generation. If migration is by females, the value of  $F_{ST} = 0.2$  corresponds to the number of immigrating females  $N_f m_f = (2 + 4x)/9x$  per generation. Note also that the haplodiploid DAM model with a sex ratio of 1:3 (males to females) gives differentiation that is independent of the sex bias among immigrants.

We close by pointing out one specific complication that arises in male-haploid social insects, where workers lay a proportion  $p$  of haploid eggs that develop into males. This leads to sex-specific reproductive values  $v_f = (2 - p)/(3 - p)$  and  $v_m = 1/(3 - p)$  (Crozier & Pamilo, 1996). This means that males get an increasing genetic influence as they contribute to the male pool of the next generation through laying workers.

## Acknowledgements

Our research has been supported by grants from the Swedish Environmental Protection Agency, Carl Trygger's Foundation and Sven and Lilly Lawski's Foundation. We thank an anonymous reviewer for valuable comments on an earlier version of this report.

## References

- AVISE, J. C. 1994. *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.
- CROW, J. F. AND KIMURA, M. 1970. *An Introduction to Population Genetics Theory*. Harper & Row, New York.
- CROZIER, R. H. AND PAMILO, P. 1996. *Evolution of Social Insect Colonies*. Oxford University Press, Oxford.
- ENNOS, R. A. 1994. Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*, **72**, 250–259.
- FELSENSTEIN, J. 1982. How can we infer geography and history from gene frequencies? *J. Theor. Biol.*, **96**, 9–20.
- GREENWOOD, P. J. 1980. Mating systems, philopatry and dispersal in birds and mammals. *Anim. Behav.*, **28**, 1140–1162.
- GREENWOOD, P. J. 1983. Mating systems and the evolutionary consequences of dispersal. In: Swingland, I. R. and Greenwood, P. J. (eds) *The Ecology of Animal Movement*, pp. 116–131. Clarendon Press, Oxford.
- LAIDLAW, H. H., JR, AND PAGE, R. E., JR, 1988. Mating designs. In: Rinderer, T. E. (ed.) *Bee Genetics and Breeding*, pp. 323–344. Academic Press, New York.
- MARUYAMA, K. AND TACHIDA, H. 1992. Genetic variability and geographical structure in partially selfing populations. *Jap. J. Genet.*, **67**, 39–51.
- OSTER, G. F., ESHEL, I. AND COHEN, D. 1977. Worker–queen conflict and the evolution of social castes. *Theor. Pop. Biol.*, **12**, 49–85.
- PAMILO, P. 1991. Evolution of colony characteristics in social insects. I. Sex allocation. *Am. Nat.*, **137**, 83–107.
- PROUT, T. 1981. A note on the island model with sex dependent migration. *Theor. Appl. Genet.*, **59**, 327–332.
- SEPPÄ, P. AND PAMILO, P. 1995. Gene flow and population viscosity in *Myrmica* ants. *Heredity*, **74**, 200–209.
- SLATKIN, M. 1985. Gene flow in natural populations. *Ann. Rev. Ecol. Syst.*, **16**, 393–430.
- SUNDSTRÖM, L. 1993. Genetic population structure and sociogenetic organization in *Formica truncorum*. *Behav. Ecol. Sociobiol.*, **33**, 345–354.
- WANG, J. 1997a. Effective size and  $F$ -statistics of subdivided populations. I. Monoecious species with partial selfing. *Genetics*, **146**, 1453–1463.
- WANG, J. 1997b. Effective size and  $F$ -statistics of subdivided populations. II. Dioecious species. *Genetics*, **146**, 1465–1474.
- WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics*, **16**, 97–159.
- WRIGHT, S. 1969. *Evolution and the Genetics of Populations*, vol. 2, *The Theory of Gene Frequencies*. University of Chicago Press, Chicago.

## Appendix

According to Crow & Kimura (1970, pp. 439–440), differentiation can be expressed with the help of first and second moments of allele frequencies,  $\mu_1 = E(x)$  and  $\mu_2 = E(x^2)$ . The moments are:

$$\begin{aligned} \mu_1^{(t+1)} &= E(x_{t+1}) = E(x_t + \delta x_t) \\ &= \mu_1^{(t)} + P_f m_f (\bar{x} - \mu_1^{(t)}) + P_m m_m (\bar{x} - \mu_1^{(t)}), \end{aligned} \quad (A1)$$

and

$$\begin{aligned} \mu_2^{(t+1)} &= E(x_{t+1}^2) = E[(x_t + \delta x_t)^2] \\ &= \mu_2^{(t)} + 2E[x_t (P_f m_f + P_m m_m) (\bar{x} - x_t)] \\ &\quad + E\left[\frac{P_f^2 x_t (1 - x_t)}{2N_f} + \frac{P_m^2 x_t (1 - x_t)}{cN_m}\right], \end{aligned} \quad (A2)$$

where  $c$  is the ploidy level of males ( $c = 2$  or  $c = 1$ ).  $P_f = P_m = 1/2$  for diploids, and  $P_f = 2/3$  and  $P_m = 1/3$  for male-haploids.

From the above equations, we get  $\Delta\mu_2^{(t)}$  and, by setting that equal to zero with increasing time, we can solve  $\sigma^2 = \mu_2^\infty - \bar{x}^2$ . Finally,  $F_{ST} = \sigma^2/\bar{x}(1 - \bar{x})$ . The result under the DBM model becomes the same as in eqns (5) and (9) for diploids and male-haploids, respectively.