Derivation of Genotypic Distribution for Three Genes under Self and Random Mating

Chen-Hung Kao and Miao-Hui Zeng

Consider the genome structure of gametes in an advanced F_t population descended from two inbred lines. For three ordered loci, A, B and C, each with two alleles ($[A,a]$, $[B,b]$ and $[C,c]$), there are eight possible gamete types, ABC, abc, AbC, aBc, ABc, abC, aBC and Abc. The union of these eight gametes can produce 36 different zygotes (27 discernible genotypes in practice). Assume that r, r_1, and r_2 denote the recombination rates between A and C, between A and B and between B and C, respectively. Let 1 and 0 represent the capital and small-letter alleles, respectively, to count the number of capital allele, so that the configurations of the eight types of gametes can be represented as $\underline{111}$, $\underline{000}$, $\underline{101}$, $\underline{010}$, $\underline{110}$, $\underline{001}$, $\underline{011}$ and $\underline{100}$, respectively, and that the genotype AA, Aa and aa (BB, Bb and bb or CC, Cc and cc) can be denoted as 2, 1 and 0, respectively. Let the eight gametic frequencies in the F_t population be $\Pr(\underline{111}) = Pr(\underline{000})=P_{1,t}$, $\Pr(\underline{101}) = Pr(\underline{010})=P_{2,t}$, $\Pr(\underline{110}) = Pr(\underline{001}) =P_{3,t}$ and $\Pr(\underline{011}) = Pr(\underline{100})=P_{4,t}$. In the F_2 population, $P_{i,t}$'s, $i = 1,2,3,4$, have simple relation with r, r_1 and r_2 as $P_{1,2} = (1-r_1)(1-r_2)/2$, $P_{2,2} = r_1 r_2/2$, $P_{3,2} = (1-r_1)r_2/2$ and $P_{4,2} = r_1(1 - r_2)/2$. However, the relation between gametic frequencies and recombination rates becomes complicated in the advanced populations after further selfing or intermating as the genomes have undergone multiple runs of meiosis. If population is mating randomly, the gametes produced by any type of zygotes (individuals) can randomly fertilize each other across different individuals. If the population is self-fertilized, the gametes of an individual are randomly mating within individual, but they are not allowed to seminate the gametes from different individuals. The procedures to obtain the transition formulas of genotypic frequencies from current to next generation under self and random mating are outlined below.

Random mating: If mating is random in a population, the unification of the eight gametes can produce 27 possible zygote genotypes, and the frequency of a zygotic genotype is equivalent to the product of the two gametic frequencies in the previous population. The genotypic distribution of a population is completely determined by the gametic distribution in the previous generation,

and the focus is on investigating how the frequencies of eight gametic types change from generation to generation.

The first key to derive the transition formulas is to make sure the all possible combinations of gametes forming each of the 27 zygotic genotypes. Among the 27 zygotic types, those (20) genotypes homozygous at two or three loci have only one gametic combination. For example, the 202 zygote is the union of two $\underline{101}$ gametes, and the 221 zygote is the combination of $\underline{111}$ and $\underline{110}$ gametes. Those (6) genotypes heterozygous at two loci have two possible gametic combinations. For example, the gametic combinations of 211 zygote can be from the union of $\underline{111}$ and $\underline{100}$ gametes or from the union of $\underline{101}$ and $\underline{110}$ gametes. The only zygote, 111, heterozygous at three loci has four possible combinations from the union of $\underline{111}$ and $\underline{000}$ gametes, $\underline{101}$ and $\underline{010}$ gametes, $\underline{001}$ and $\underline{110}$ gametes, or union of $\underline{011}$ and $\underline{100}$ gametes. The second key is to infer the possible gametic types produced by each of the possible combinations (genotypes) and to calculate the produced proportions. Table 1 shows the possible types and proportions of gametes produced by each possible genotype.

With the possible gametic combinations of a zygote and the possible types and proportions of gametes a zygote can produced, the transition equations of gametic frequencies from the current generation to next generation can be obtained. The transition equation of the frequency, $P_{1,t+1}$, for $\underline{11}$ ($\underline{000}$) gamete in terms of gametic frequencies in the previous generation, $P_{i,t}$'s, can be obtained as

$$
\begin{aligned}
P_{1,t+1} \quad = \quad & [(1-r_1)(1-r_2)+1]P_{1,t}^2 + r_1 r_2 P_{2,t}^2 + [(1-r_1)r_2]P_{3,t}^2 + [r_1(1-r_2)]P_{4,t}^2 \\
& + [1+(1-r_1)(1-r_2)+r_1 r_2]P_{1,t}P_{2,t} + (2-r_1)P_{1,t}P_{3,t} + (2-r_2)P_{1,t}P_{4,t} \\
& + r_2 P_{2,t}P_{3,t} + r_1 P_{2,t}P_{4,t} + [r_1(1-r_2)+r_2(1-r_1)]P_{3,t}P_{4,t}, \quad\quad\quad (1)
\end{aligned}
$$

The above equation is obtained as follows. For ($\underline{111}$) gamete in F_{t+1} population, it can be produced by eight different parental zygotes 222, 212, 221, 211, 122, 112, 121 and 111. The proportions of $\underline{111}$ gamete produced by these eight zygotes are calculated as follows. For 222 zygote, the proportion to produce $\underline{111}$ gamete is one. For 212, 221 and 122 zygotes, the proportions are all 0.5. For 121 zygote, the two gametes forming this zygote can be $\underline{111}$ and $\underline{010}$, or $\underline{110}$ and $\underline{011}$. If 121 zygote is from union of $\underline{111}$ and $\underline{010}$ gametes, the proportion to produce $\underline{111}$ gamete in the next generation is $(1-r_1)(1-r_2)+r_1 r_2$. If it is from union of $\underline{111}$ and $\underline{010}$ gametes, the proportion in the next generation is $r_1(1-r_2)+r_2(1-r_1)$. For 211 (112) zygote, the two

2

gametes forming this zygote can be $\underline{111}$ and $\underline{100}$ ($\underline{111}$ and $\underline{001}$), or $\underline{101}$ and $\underline{110}$ ($\underline{101}$ and $\underline{011}$). For zygote 211 (112) from union of $\underline{111}$ and $\underline{100}$ ($\underline{111}$ and $\underline{001}$) gametes, the proportion to produce $\underline{111}$ gamete in the next generation is $(1 - r_2)$ $((1 - r_1))$. For zygote 211 (112) from union of $\underline{101}$ and $\underline{110}$ ($\underline{101}$ and $\underline{011}$) gametes, the proportion to produce $\underline{111}$ gamete is r_2 (r_1). For 111 zygote, there are four kinds of gamete unions, $\underline{111}$ with $\underline{000}$, $\underline{101}$ with $\underline{010}$, $\underline{001}$ with $\underline{110}$, and $\underline{011}$ with $\underline{100}$. The proportions to produce $\underline{111}$ gamete by 111 zygote from the four unions are $(1 - r_1)(1 - r_2)$, $r_1 r_2$, $(1 - r_1)r_2$ and $r_1(1 - r_2)$, respectively. As these eight zygotes capable to produce $\underline{111}$ gamete have frequencies $P_{1,t}^2$, $2P_{1,t}P_{2,t}$, $2P_{1,t}P_{3,t}$, $2(P_{1,t}P_{4,t}+P_{2,t}P_{3,t})$, $2P_{1,t}P_{4,t}$, $2(P_{1,t}P_{3,t}+P_{2,t}P_{4,t})$, $2(P_{1,t}P_{2,t}+P_{3,t}P_{4,t})$, $P_{1,t}^2+P_{2,t}^2+P_{3,t}^2+P_{4,t}^2$, respectively, in the F_t population, the frequency of $\underline{111}$ gamete in the F_{t+1} population can be obtained as the sum of the products of zygote frequencies in the previous generation and producing proportions as in Equation (1). By a similar argument, the iteration equations of the remaining three gametic frequencies $P_{2,t+1}$, $P_{3,t+1}$, and $P_{4,t+1}$ can be obtained as

$$
\begin{aligned}
P_{2,t} = \ & r_1 r_2 P_{1,t-1}^2 + [(1 - r_1)(1 - r_2) + 1]P_{2,t-1}^2 + r_1(1 - r_2)P_{3,t-1}^2 \\
& + (1 - r_1)r_2 P_{4,t-1}^2 + [1 + (1 - r_1)(1 - r_2) + r_1 r_2]P_{1,t-1}P_{2,t-1} \\
& + r_1 P_{1,t-1}P_{3,t-1} + r_2 P_{1,t-1}P_{4,t-1} + (2 - r_2)P_{2,t-1}P_{3,t-1} \\
& + (2 - r_1)P_{2,t-1}P_{4,t-1} + [r_1(1 - r_2) + r_2(1 - r_1)]P_{3,t-1}P_{4,t-1},
\end{aligned} \tag{2}
$$

$$
\begin{aligned}
P_{3,t} = \ & (1 - r_1)r_2 P_{1,t-1}^2 + r_1(1 - r_2)P_{2,t-1}^2 + [(1 - r_1)(1 - r_2) + 1]P_{3,t-1}^2 \\
& + r_1 r_2 P_{4,t-1}^2 + [r_1(1 - r_2) + (1 - r_1)r_2]P_{1,t-1}P_{2,t-1} \\
& + (2 - r_1)P_{1,t-1}P_{3,t-1} + r_2 P_{1,t-1}P_{4,t-1} + (2 - r_2)P_{2,t-1}P_{3,t-1} \\
& + r_1 P_{2,t-1}P_{4,t-1} + [(1 - r_1)(1 - r_2) + r_1 r_2 + 1]P_{3,t-1}P_{4,t-1},
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
P_{4,t} = \ & r_1(1 - r_2)P_{1,t-1}^2 + (1 - r_1)r_2 P_{2,t-1}^2 + r_1 r_2 P_{3,t-1}^2 \\
& + [(1 - r_1)(1 - r_2) + 1]P_{4,t-1}^2 + [r_1(1 - r_2) + (1 - r_1)r_2]P_{1,t-1}P_{2,t-1} \\
& + r_1 P_{1,t-1}P_{3,t-1} + (2 - r_2)P_{1,t-1}P_{4,t-1} + r_2 P_{2,t-1}P_{3,t-1} \\
& + (2 - r_1)P_{2,t-1}P_{4,t-1} + [(1 - r_1)(1 - r_2) + r_1 r_2 + 1]P_{3,t-1}P_{4,t-1},
\end{aligned} \tag{4}
$$

With the above iteration equations, it is sufficient to obtain the gametic frequencies in any generation in terms of r_1, r_2 and r. Once the eight gametic frequencies are obtained in a specific F_t population, the actual frequencies of the 27 genotypes can be derived by random union of the gametes. The final proportions for the eight gamete types are $P_{1,\infty} = P_{2,\infty} = P_{3,\infty} = P_{4,\infty} = 1/8$.

In programming the transition equations, directly applying equations (1) to (4) may cause large rounding errors for large t, and rearrangement of the equations is needed to ensure stable convergence (see website http://www.stat.sinica.edu.tw/∼chkao). In general, their convergence is quicker for larger r.

Selfing: If population is self-fertilized, the gametes are restricted to fertilize the gametes from the same parents and not transmitted between different parents. In a population, the genotypic frequencies in t generation are determined by the genotypic frequencies in the previous $t - 1$ generation. The focus is therefore on deriving the frequencies of genotypes, rather than on those of gametes from generation to generation.

To obtain the iteration equations of each genotypes, it is required to identify the coupling and repulsion phases for the double and triple heterozygous zygotes. If the phases of zygote genotypes are discernible, there are 36 different zygotes. Among the 36 zygotic genotypes, the four zygotic types heterozygous at three loci can produce the all eight gamete types and propagate the all 36 zygotic types, and remaining zygotes can produce only some (four, two and one) of the gametic types and propagate part of the 36 zygotic types in the progeny population. In sum, the zygotes producing one, two, four and eight gamete types can propagate one, three, ten and 36 zygotic types in the progeny populations. Let the 36 zygotic genotypes occur in a population with the following proportions:

$$
\begin{array}{llllllll}
A_n & \frac{111}{111} & \frac{000}{000} & B_n & \frac{111}{101} & \frac{000}{010} & C_n & \frac{101}{101} & \frac{010}{010} \\[6pt]
D_n & \frac{111}{110} & \frac{000}{001} & E_n & \frac{111}{100} & \frac{000}{011} & F_n & \frac{101}{110} & \frac{010}{001} \\[6pt]
G_n & \frac{101}{100} & \frac{010}{011} & H_n & \frac{110}{110} & \frac{001}{001} & I_n & \frac{110}{100} & \frac{001}{011} \\[6pt]
J_n & \frac{100}{100} & \frac{011}{011} & K_n & \frac{111}{011} & \frac{000}{100} & L_n & \frac{111}{001} & \frac{000}{110} \\[6pt]
M_n & \frac{101}{011} & \frac{010}{100} & N_n & \frac{101}{001} & \frac{010}{110} & S_n & \frac{111}{010} & \frac{000}{101} \\[6pt]
T_n & \frac{110}{011} & \frac{001}{100} & U_n & \frac{111}{000} & & V_n & \frac{101}{010} & \\[6pt]
W_n & \frac{001}{110} & & Z_n & \frac{100}{011} & & &
\end{array}
$$

Among the 36 zygotic genotypes, eight zygotes, $\frac{111}{111}, \frac{000}{000}, \frac{101}{101}, \frac{010}{010}, \frac{110}{110}, \frac{001}{001}, \frac{100}{100}$ and $\frac{011}{011}$, homozygous at three loci can produce only one gametic type (see Table 1). Twelve zygotes $\frac{111}{101}, \frac{000}{010}, \frac{111}{110}, \frac{000}{001}, \frac{101}{100},$ $\frac{010}{011}, \frac{110}{100}, \frac{001}{011}, \frac{111}{011}, \frac{000}{100}, \frac{101}{001}$ and $\frac{010}{110}$ heterozygous at one locus are capable of producing two gametic types. Fourteen zygotes $\frac{111}{100}, \frac{000}{011}, \frac{101}{110}, \frac{010}{001}, \frac{111}{001}, \frac{000}{110}, \frac{101}{011}, \frac{010}{100}, \frac{111}{010}, \frac{000}{101}, \frac{110}{011}$ and $\frac{001}{100}$ heterozygous at two loci can produce four gametic types. Four zygotes, $\frac{111}{000}, \frac{101}{010}, \frac{001}{110}, \frac{100}{011}$ heterozygous at three loci

can produce the all eight possible gametic types. We will have the following results for $\frac{111}{111}$ zygote:

$$A_{n+1} = A_n + \frac{1}{4}B_n + \frac{1}{4}D_n + \frac{(1-r_2)^2}{4}E_n + \frac{r_2^2}{4}F_n + \frac{1}{4}K_n + \frac{(1-r_1)^2}{4}L_n + \frac{r_1^2}{4}M_n$$
$$+ \frac{[(1-r_1)(1-r_2)+r_1r_2]^2}{4}S_n + \frac{[r_1(1-r_2)+r_2(1-r_1)]^2}{4}T_n$$
$$+ \frac{[(1-r_1)(1-r_2)]^2}{4}U_n + \frac{r_1^2r_2^2}{4}V_n + \frac{[(1-r_1)r_2]^2}{4}W_n + \frac{[r_1(1-r_2)]^2}{4}Z_n \qquad (5)$$

This equation is derived as follows. The $\frac{111}{111}$ ($\frac{000}{000}$) zygote is from the union of two $\underline{111}(\underline{000})$ gametes. By Table 1 and above description, fourteen zygotes are capable of contributing $\underline{111}(\underline{000})$ gametes with different proportions. Under selfing, the frequency of the $\frac{111}{111}$ ($\frac{000}{000}$) genotype in the progeny population is equivalent to the sum of the frequencies of $\frac{111}{111}$ progeny produced by each of the the fourteen possible parental zygotes. The proportion of contribution of A_n to A_{n+1} is 100%. The proportions of the contribution of B_n, D_n and K_n (zygotes heterozygous at one locus) to A_{n+1} is $(1/2)^2$. The proportions of the contribution of six zygotes heterozygous at double loci (E_n, F_n, L_n, M_n, S_n and T_n) are $(1-r_2)^2/4$, $r_2^2/4$, $(1-r_1)^2/4$, $r_1^2/4$, $[(1-r_1)(1-r_2)+r_1r_2]^2/4$ and $[r_1(1-r_2)+r_2(1-r_1)]^2/4$, respectively. The proportions of contribution of the four genotypes heterozygous at three loci are $[(1-r_1)(1-r_2)]^2/4$, $(r_1r_2)^2/4$, $[(1-r_1)r_2]^2/4$ and $[r_1(1-r_2)]^2/4$ for four different phases. The iteration equations of other zygotes can be similarly derived, and they are as follows.

$$B_{n+1} = \frac{1}{2}B_n + \frac{r_2(1-r_2)}{2}E_n + \frac{r_2(1-r_2)}{2}F_n + \frac{r_1(1-r_1)}{2}L_n + \frac{r_1(1-r_1)}{2}M_n$$
$$+ \frac{r_1r_2(1-r_1)(1-r_2)}{2}U_n + \frac{r_1r_2(1-r_1)(1-r_2)}{2}V_n$$
$$+ \frac{r_1r_2(1-r_1)(1-r_2)}{2}W_n + \frac{r_1r_2(1-r_1)(1-r_2)}{2}Z_n \qquad (6)$$

$$C_{n+1} = \frac{1}{4}B_n + C_n + \frac{r_2^2}{4}E_n + \frac{(1-r_2)^2}{4}F_n + \frac{1}{4}G_n + \frac{r_1^2}{4}L_n + \frac{(1-r_1)^2}{4}M_n + \frac{1}{4}N_n$$
$$+ \frac{[(1-r_1)(1-r_2)+r_1r_2]^2}{4}S_n + \frac{[r_1(1-r_2)+r_2(1-r_1)]^2}{4}T_n$$
$$+ \frac{r_1^2r_2^2}{4}U_n + \frac{[(1-r_1)(1-r_2)]^2}{4}V_n + \frac{[r_1(1-r_2)]^2}{4}W_n + \frac{[(1-r_1)r_2]^2}{4}Z_n \qquad (7)$$

$$D_{n+1} = \frac{1}{2}D_n + \frac{r_2(1-r_2)}{2}E_n + \frac{r_2(1-r_2)}{2}F_n$$
$$+ \frac{[(1-r_1)(1-r_2)+r_1r_2][r_1(1-r_2)+r_2(1-r_1)]}{2}S_n$$
$$+ \frac{[(1-r_1)(1-r_2)+r_1r_2][r_1(1-r_2)+r_2(1-r_1)]}{2}T_n$$
$$+ \frac{(1-r_1)^2[r_2(1-r_2)]}{2}U_n + \frac{r_1^2[r_2(1-r_2)]}{2}V_n$$

$$+\frac{(1-r_1)^2[r_2(1-r_2)]}{2}W_n + \frac{r_1^2[r_2(1-r_2)]}{2}Z_n \tag{8}$$

$$E_{n+1} = \frac{(1-r_2)^2}{2}E_n + \frac{r_2^2}{2}F_n + \frac{r_1(1-r_1)(1-r_2)^2}{2}U_n + \frac{r_1(1-r_1)r_2^2}{2}V_n$$
$$+\frac{r_1(1-r_1)r_2^2}{2}W_n + \frac{r_1(1-r_1)(1-r_2)^2}{2}Z_n \tag{9}$$

$$F_{n+1} = \frac{r_2^2}{2}E_n + \frac{(1-r_2)^2}{2}F_n + \frac{r_1(1-r_1)r_2^2}{2}U_n + \frac{r_1(1-r_1)(1-r_2)^2}{2}V_n$$
$$+\frac{r_1(1-r_1)(1-r_2)^2}{2}W_n + \frac{r_1(1-r_1)r_2^2}{2}Z_n \tag{10}$$

$$G_{n+1} = \frac{r_2(1-r_2)}{2}E_n + \frac{r_2(1-r_2)}{2}F_n + \frac{1}{2}G_n$$
$$+\frac{[(1-r_1)(1-r_2)+r_1r_2][r_1(1-r_2)+r_2(1-r_1)]}{2}S_n$$
$$+\frac{[(1-r_1)(1-r_2)+r_1r_2][r_1(1-r_2)+r_2(1-r_1)]}{2}T_n$$
$$+\frac{r_1^2r_2(1-r_2)}{2}U_n + \frac{(1-r_1)^2r_2(1-r_2)}{2}V_n$$
$$+\frac{r_1^2r_2(1-r_2)}{2}W_n + \frac{(1-r_1)^2r_2(1-r_2)}{2}Z_n \tag{11}$$

$$H_{n+1} = \frac{1}{4}D_n + \frac{r_2^2}{4}E_n + \frac{(1-r_2)^2}{4}F_n + H_n + \frac{1}{4}I_n + \frac{(1-r_1)^2}{4}L_n + \frac{r_1^2}{4}M_n + \frac{1}{4}N_n$$
$$+\frac{[r_1(1-r_2)+(1-r_1)r_2]^2}{4}S_n + \frac{[(1-r_1)(1-r_2)+r_1r_2]^2}{4}T_n$$
$$+\frac{[(1-r_1)r_2]^2}{4}U_n + \frac{[r_1(1-r_2)]^2}{4}V_n + \frac{[(1-r_1)(1-r_2)]^2}{4}W_n + \frac{r_1^2r_2^2}{4}Z_n \tag{12}$$

$$I_{n+1} = \frac{r_2(1-r_2)}{2}E_n + \frac{r_2(1-r_2)}{2}F_n + \frac{1}{2}I_n + \frac{r_1(1-r_1)}{2}L_n + \frac{r_1(1-r_1)}{2}M_n$$
$$+\frac{r_1(1-r_1)r_2(1-r_2)}{2}U_n + \frac{r_1(1-r_1)r_2(1-r_2)}{2}V_n$$
$$+\frac{r_1(1-r_1)r_2(1-r_2)}{2}W_n + \frac{r_1(1-r_1)r_2(1-r_2)}{2}Z_n \tag{13}$$

$$J_{n+1} = \frac{(1-r_2)^2}{4}E_n + \frac{r_2^2}{4}F_n + \frac{1}{4}G_n + \frac{1}{4}I_n + J_n + \frac{1}{4}K_n + \frac{r_1^2}{4}L_n + \frac{(1-r_1)^2}{4}M_n$$
$$+\frac{[r_1(1-r_2)+r_2(1-r_1)]^2}{4}S_n + \frac{[(1-r_1)(1-r_2)+r_1r_2]^2}{4}T_n$$
$$+\frac{[r_1(1-r_2)]^2}{4}U_n + \frac{[(1-r_1)r_2]^2}{4}V_n + \frac{r_1^2r_2^2}{4}W_n + \frac{[(1-r_1)(1-r_2)]^2}{4}Z_n \tag{14}$$

$$K_{n+1} = \frac{1}{2}K_n + \frac{r_1(1-r_1)}{2}L_n + \frac{r_1(1-r_1)}{2}M_n$$
$$+\frac{[(1-r_1)(1-r_2)+r_1r_2][r_1(1-r_2)+r_2(1-r_1)]}{2}S_n$$
$$+\frac{[(1-r_1)(1-r_2)+r_1r_2][r_1(1-r_2)+r_2(1-r_1)]}{2}T_n$$

$$+\frac{r_1(1-r_1)(1-r_2)^2}{2}U_n + \frac{r_1(1-r_1)r_2^2}{2}V_n$$

$$+\frac{r_1(1-r_1)r_2^2}{2}W_n + \frac{r_1(1-r_1)(1-r_2)^2}{2}Z_n \tag{15}$$

$$L_{n+1} = \frac{(1-r_1)^2}{2}L_n + \frac{r_1^2}{2}M_n + \frac{(1-r_1)^2 r_2(1-r_2)}{2}U_n + \frac{r_1^2 r_2(1-r_2)}{2}V_n$$

$$+\frac{(1-r_1)^2 r_2(1-r_2)}{2}W_n + \frac{r_1^2 r_2(1-r_2)}{2}Z_n \tag{16}$$

$$M_{n+1} = \frac{r_1^2}{2}L_n + \frac{(1-r_1)^2}{2}M_n + \frac{r_1^2 r_2(1-r_2)}{2}U_n + \frac{(1-r_1)^2 r_2(1-r_2)}{2}V_n$$

$$+\frac{r_1^2 r_2(1-r_2)}{2}W_n + \frac{(1-r_1)^2 r_2(1-r_2)}{2}Z_n \tag{17}$$

$$N_{n+1} = \frac{r_1(1-r_1)}{2}L_n + \frac{r_1(1-r_1)}{2}M_n + \frac{1}{2}N_n$$

$$+\frac{[(1-r_1)(1-r_2)+r_1 r_2][r_1(1-r_2)+r_2(1-r_1)]}{2}S_n$$

$$+\frac{[(1-r_1)(1-r_2)+r_1 r_2][r_1(1-r_2)+r_2(1-r_1)]}{2}T_n$$

$$+\frac{r_1(1-r_1)r_2^2}{2}U_n + \frac{r_1(1-r_1)(1-r_2)^2}{2}V_n$$

$$+\frac{r_1(1-r_1)(1-r_2)^2}{2}W_n + \frac{r_1(1-r_1)r_2^2}{2}Z_n \tag{18}$$

$$S_{n+1} = \frac{[(1-r_1)(1-r_2)+r_1 r_2]^2}{2}S_n + \frac{[r_1(1-r_2)+r_2(1-r_1)]^2}{2}T_n$$

$$+\frac{r_1(1-r_1)r_2(1-r_2)}{2}U_n + \frac{r_1(1-r_1)r_2(1-r_2)}{2}V_n$$

$$+\frac{r_1(1-r_1)r_2(1-r_2)}{2}W_n + \frac{r_1(1-r_1)r_2(1-r_2)}{2}Z_n \tag{19}$$

$$T_{n+1} = \frac{[r_1(1-r_2)+r_2(1-r_1)]^2}{2}S_n + \frac{[(1-r_1)(1-r_2)+r_1 r_2]^2}{2}T_n$$

$$+\frac{r_1(1-r_1)r_2(1-r_2)}{2}U_n + \frac{r_1(1-r_1)r_2(1-r_2)}{2}V_n$$

$$+\frac{r_1(1-r_1)r_2(1-r_2)}{2}W_n + \frac{r_1(1-r_1)r_2(1-r_2)}{2}Z_n \tag{20}$$

$$U_{n+1} = \frac{[(1-r_1)(1-r_2)]^2}{2}U_n + \frac{r_1^2 r_2^2}{2}V_n + \frac{[(1-r_1)r_2]^2}{2}W_n + \frac{[r_1(1-r_2)]^2}{2}Z_n \tag{21}$$

$$V_{n+1} = \frac{r_1^2 r_2^2}{2}U_n + \frac{[(1-r_1)(1-r_2)]^2}{2}V_n + \frac{[r_1(1-r_2)]^2}{2}W_n + \frac{[(1-r_1)r_2]^2}{2}Z_n \tag{22}$$

$$W_{n+1} = \frac{[(1-r_1)r_2]^2}{2}U_n + \frac{[r_1(1-r_2)]^2}{2}V_n + \frac{[(1-r_1)(1-r_2)]^2}{2}W_n + \frac{[r_1 r_2]^2}{2}Z_n \tag{23}$$

$$Z_{n+1} = \frac{[r_1(1-r_2)]^2}{2}U_n + \frac{[(1-r_1)r_2]^2}{2}V_n + \frac{[r_1 r_2]^2}{2}W_n + \frac{[(1-r_1)(1-r_2)]^2}{2}Z_n \tag{24}$$

The SAS program for these transition equations are placed at http://www.stat.sinica.edu.tw/~chkao. The actual frequencies of different zygotic types in each generation can be calculated from equa-

7

tion (5) to (24). Clearly, the final frequencies of zygotes with heterozygous locus, B_∞, D_∞, E_∞, F_∞, G_∞, I_∞, K_∞, L_∞, M_∞, N_∞, S_∞, T_∞, U_∞, V_∞, W_∞ and Z_∞, are zero. The final proportions for the eight types of zygotes homozygous at three loci, A_∞, C_∞, H_∞ and J_∞, are

$$
\begin{aligned}
A_\infty &= \frac{1}{4}\left(2 - \frac{2r_1}{1+2r_1} - \frac{2r_2}{1+2r_2} - \frac{2r}{1+2r}\right), \\
C_\infty &= \frac{1}{4}\left(-\frac{2r_1}{1+2r_1} + \frac{2r_2}{1+2r_2} + \frac{2r}{1+2r}\right), \\
H_\infty &= \frac{1}{4}\left(\frac{2r_1}{1+2r_1} - \frac{2r_2}{1+2r_2} + \frac{2r}{1+2r}\right), \\
J_\infty &= \frac{1}{4}\left(\frac{2r_1}{1+2r_1} + \frac{2r_2}{1+2r_2} - \frac{2r}{1+2r}\right)
\end{aligned}
$$

according to Geiringer (1944) and Martin and Hospital (2006). For example, if $r_1 = 0.1$ and $r_2 = 0.1$, the final values of A, C, G and J are 0.35049, 0.01716, 0.06618 and 0.06618, respectively.

Random mating after selfing: If the individuals are selfed for t generations and then randomly intermated for j $(j > 0)$ generations to create new progeny (RIX) populations, the genotypic frequencies can be obtained by first deriving the genotypic frequencies in the F_t generation using the iteration equations (5) to (24), and then calculating the gametic proportions, $P_{i,t}$, $i = 1, 2, 3, 4$, for plugging in the iteration equations of random mating to obtain the new gametic frequencies in RIX $F_{t,j}$ population. For example, the proportions of the 111 gamete in the F_t populations after t generations of selfing are calculated as

$$
\begin{aligned}
P_{1,t+1} &= A_t + \frac{1}{2}B_t + \frac{1}{2}D_t\frac{1-r_2}{2}E_t + \frac{r_2}{2}F_t + \frac{1}{2}K_t + \frac{1-r_1}{2}L_t + \frac{r_1}{2}M_t \\
&\quad + \frac{(1-r_1)(1-r_2) + r_1 r_2}{2}S_t + \frac{r_1(1-r_2) + r_2(1-r_1)}{2}T_t \\
&\quad + \frac{(1-r_1)(1-r_2)}{2}U_t + \frac{r_1 r_2}{2}V_t + \frac{(1-r_1)r_2}{2}W_t + \frac{r_1(1-r_2)}{2}Z_t.
\end{aligned}
$$

Similarly, $P_{2,t}$, $P_{3,t}$ and $P_{4,t}$ can be obtained. These gametic frequencies are then used in the iteration equations to produce the gametic frequencies after j generations of random mating, and in turn to obtain genotypic distribution of the RIX $F_{t,j}$ populations.

Selfing after random mating: If the individuals in the AI F_t population are selfed for j generations to produce the IRI $F_{t,j}$ populations, the genotypic distribution in the IRI $F_{t,j}$ populations can be obtained by first deriving the eight gametic frequencies, $P_{i,t}$, $i = 1, 2, 3, 4$, in the AI F_t populations, and further to obtain the genotypic frequencies bu using

$$A_t = P_{1,t}^2 \qquad B_t = 2P_{1,t}P_{2,t} \qquad C_t = P_{2,t}^2 \qquad D_t = 2P_{1,t}P_{3,t}$$

$$E_t = 2P_{1,t}P_{4,t} \qquad F_t = 2P_{2,t}P_{3,t} \qquad G_t = 2P_{2,t}P_{4,t} \qquad H_t = P_{3,t}^2$$

$$I_t = 2P_{3,t}P_{4,t} \qquad J_t = P_{4,t}^2 \qquad K_t = 2P_{1,t}P_{4,t} \qquad L_t = 2P_{1,t}P_{3,t}$$

$$M_t = 2P_{2,t}P_{4,t} \qquad N_t = 2P_{2,t}P_{3,t} \qquad S_t = 2P_{1,t}P_{2,t} \qquad T_t = 2P_{3,t}P_{4,t}$$

$$U_t = 2P_{1,t}^2 \qquad V_t = 2P_{2,t}^2 \qquad W_t = 2P_{3,t}^2 \qquad Z_t = 2P_{4,t}^2,$$

These genotypic frequencies in AI F_t population are used in equations (5) to (24) to obtain the genotypic frequencies of the IRI $F_{t,j}$ populations after t generations of selfing.

LITERATURE CITED

Geiringer, H., 1944 On the probability theory of linkage in Mendelian heredity. *The Annals of Mathematical Statistics* **15**: 25- 57.

Martin, O. C. and F. Hospital, 2006 Two-and three-locus tests for linkage analysis using recombinant inbred lines. *Genetics* **173**: 451- 459.

Table 1: the types and proportions of gametes produced by each possible zygotic genotype

Zygote type	Types of gametes produced							
	111	101	110	011	100	001	010	000
$\frac{111}{111}$	1							(1)
$\frac{111}{101}$	$\frac{1}{2}$	$\frac{1}{2}$					$\left(\frac{1}{2}\right)$	$\left(\frac{1}{2}\right)$
$\frac{101}{101}$		1					(1)	
$\frac{111}{110}$	$\frac{1}{2}$		$\frac{1}{2}$			$\left(\frac{1}{2}\right)$		$\left(\frac{1}{2}\right)$
$\frac{111}{100}$	$\frac{1-r_2}{2}$	$\frac{r_2}{2}$	$\frac{r_2}{2}$	$\left(\frac{1-r_2}{2}\right)$	$\frac{1-r_2}{2}$	$\left(\frac{r_2}{2}\right)$	$\left(\frac{r_2}{2}\right)$	$\left(\frac{1-r_2}{2}\right)$
$\frac{101}{110}$	$\frac{r_2}{2}$	$\frac{1-r_2}{2}$	$\frac{1-r_2}{2}$	$\left(\frac{r_2}{2}\right)$	$\frac{r_2}{2}$	$\left(\frac{1-r_2}{2}\right)$	$\left(\frac{1-r_2}{2}\right)$	$\left(\frac{r_2}{2}\right)$
$\frac{101}{100}$		$\frac{1}{2}$		$\left(\frac{1}{2}\right)$	$\frac{1}{2}$		$\left(\frac{1}{2}\right)$	
$\frac{110}{110}$			1			(1)		
$\frac{110}{100}$			$\frac{1}{2}$	$\left(\frac{1}{2}\right)$	$\frac{1}{2}$	$\left(\frac{1}{2}\right)$		
$\frac{100}{100}$				(1)	1			
$\frac{111}{011}$	$\frac{1}{2}$			$\frac{1}{2}$	$\left(\frac{1}{2}\right)$			$\left(\frac{1}{2}\right)$
$\frac{111}{001}$	$\frac{1-r_1}{2}$	$\frac{r_1}{2}$	$\left(\frac{1-r_1}{2}\right)$	$\frac{r_1}{2}$	$\left(\frac{r_1}{2}\right)$	$\frac{1-r_1}{2}$	$\left(\frac{r_1}{2}\right)$	$\left(\frac{1-r_1}{2}\right)$
$\frac{101}{011}$	$\frac{r_1}{2}$	$\frac{1-r_1}{2}$	$\left(\frac{r_1}{2}\right)$	$\frac{1-r_1}{2}$	$\left(\frac{1-r_1}{2}\right)$	$\frac{r_1}{2}$	$\left(\frac{1-r_1}{2}\right)$	$\left(\frac{r_1}{2}\right)$
$\frac{101}{001}$		$\frac{1}{2}$	$\left(\frac{1}{2}\right)$			$\frac{1}{2}$	$\left(\frac{1}{2}\right)$	
$\frac{111}{010}$	$\frac{1-r}{2}$	$\left(\frac{1-r}{2}\right)$	$\frac{r}{2}$	$\frac{r}{2}$	$\left(\frac{r}{2}\right)$	$\left(\frac{r}{2}\right)$	$\frac{1-r}{2}$	$\left(\frac{1-r}{2}\right)$
$\frac{110}{011}$	$\frac{r}{2}$	$\left(\frac{r}{2}\right)$	$\frac{1-r}{2}$	$\frac{1-r}{2}$	$\left(\frac{1-r}{2}\right)$	$\left(\frac{1-r}{2}\right)$	$\frac{r}{2}$	$\left(\frac{r}{2}\right)$
$\frac{111}{000}$	$\frac{(1-r_1)(1-r_2)}{2}$	$\frac{r_1 r_2}{2}$	$\frac{(1-r_1)r_2}{2}$	$\frac{r_1(1-r_2)}{2}$	$\frac{r_1(1-r_2)}{2}$	$\frac{(1-r_1)r_2}{2}$	$\frac{r_1 r_2}{2}$	$\frac{(1-r_1)(1}{2}$
$\frac{101}{010}$	$\frac{r_1 r_2}{2}$	$\frac{(1-r_1)(1-r_2)}{2}$	$\frac{r_1(1-r_2)}{2}$	$\frac{(1-r_1)r_2}{2}$	$\frac{(1-r_1)r_2}{2}$	$\frac{r_1(1-r_2)}{2}$	$\frac{(1-r_1)(1-r_2)}{2}$	$\frac{r_1 r_2}{2}$
$\frac{001}{110}$	$\frac{(1-r_1)r_2}{2}$	$\frac{r_1(1-r_2)}{2}$	$\frac{(1-r_1)(1-r_2)}{2}$	$\frac{r_1 r_2}{2}$	$\frac{r_1 r_2}{2}$	$\frac{(1-r_1)(1-r_2)}{2}$	$\frac{r_1(1-r_2)}{2}$	$\frac{(1-r_1)}{2}$
$\frac{100}{011}$	$\frac{r_1(1-r_2)}{2}$	$\frac{(1-r_1)r_2}{2}$	$\frac{r_1 r_2}{2}$	$\frac{(1-r_1)(1-r_2)}{2}$	$\frac{(1-r_1)(1-r_2)}{2}$	$\frac{r_1 r_2}{2}$	$\frac{(1-r_1)r_2}{2}$	$\frac{r_1(1-}{2}$