SIB MATING WITH TWO LINKED LOCI¹

C. CLARK COCKERHAM AND B. S. WEIR²

Department of Experimental Statistics, North Carolina State University, Raleigh, N.C.

Received February 5, 1968

SIB mating with one locus has been studied extensively by several authors. WRIGHT (1921) in terms of the correlation between uniting gametes, HALDANE (1937) using a generation matrix for the six possible mating types, and MALÉCOT (1948) using probability arguments, all produced the same recurrence relation of a measure of the degree of homozygosity, the inbreeding coefficient. FISHER (1949) continued the generation matrix approach and found the genotypic array after n generations of sib mating.

When several linked loci are considered, the problem of establishing a measure of the amount of inbreeding which lends itself readily to evaluation in all generations is more difficult. HALDANE and WADDINGTON (1931) reckoned with one hundred mating types for two alleles at each of two loci, and found the final frequency of crossover zygotes. HALDANE (1949) extended the inbreeding coefficient to two loci, and gave its value for the offspring of full sibs from unrelated, outbred parents (the first generation of inbreeding). BENNETT (1954) returned to the generation matrix approach, restricting himself to two alleles and the nine mating types that gave offspring heterozygous at both loci, and listed numerical values of the dominant root of the matrix. SCHNELL (1961) introduced a general function of inbreeding for any number of linked loci and gave its value for only the first generation of inbreeding for sib mating. SHIKATA (1962) also generalized the inbreeding coefficient, and gave a recurrence relation for it for the two locus sib mating scheme (SHIKATA 1965a). However, because of limited applicability of this method (SHIKATA 1965b) the formula is incorrect.

The effects of linkage on inbreeding in selfing populations are now known. NARAIN (1965), using probability measures (only two types required), has solved recurrence relations for the inbreeding coefficient for an arbitrary number of linked loci. However, the analysis of sib mating, the next less intense form of inbreeding, has not been completed.

In this paper are defined additional measures of the probability of genes at two loci being identical by descent. Transition equations for the set of measures are given for sib mating and a recurrence relation is found for the panmictic function.

¹ Paper Number 2770 of the Journal Series of the North Carolina State University Agricultural Experiment Station, Raleigh, N. C. This investigation was supported in part by Public Health Service Research Grant GM 11546 from the Division of General Medical Sciences. The computing services for this investigation were provided by NIH Grant No. FR-00011.

² On leave from the Applied Mathematics Division of DSIR, Wellington, New Zealand.

DEFINITION OF MEASURES (GENERAL)

We consider two linked autosomal loci, the first having alleles $a_1, a_2, \dots$ and the second having alleles $b_1, b_2, \dots$. For various sets (a_r, b_t, a_s, b_u) of two genes for each locus we define functions of the form $X_{ij}(a_r, b_t, a_s, b_u)$, $(i, j = 0, 1)$. If these genes are located on gametes that form generation t , we write the functions as X^t_{ij} . The definitions of each component of the functions are:

$$\begin{aligned} X^t_{11} &= \text{Prob } [a_r \equiv a_s, b_t \equiv b_u] \\ X^t_{10} &= \text{Prob } [a_r \equiv a_s, b_t \not\equiv b_u] \\ X^t_{01} &= \text{Prob } [a_r \not\equiv a_s, b_t \equiv b_u] \\ X^t_{00} &= \text{Prob } [a_r \not\equiv a_s, b_t \not\equiv b_u] \end{aligned} \quad (1)$$

so that

$$\begin{aligned} X^t_{1.} &= X^t_{11} + X^t_{10} = \text{Prob } [a_r \equiv a_s] \\ X^t_{.1} &= X^t_{11} + X^t_{01} = \text{Prob } [b_t \equiv b_u] \\ X^t_{0.} &= X^t_{01} + X^t_{00} = \text{Prob } [a_r \not\equiv a_s] \\ X^t_{.0} &= X^t_{01} + X^t_{00} = \text{Prob } [b_t \not\equiv b_u] \\ X^t_{..} &= X^t_{11} + X^t_{10} + X^t_{01} + X^t_{00} = 1 \end{aligned} \quad (2)$$

where the identity relation $a_r \equiv a_s$ means that a_r and a_s are genes identical by descent.

DEFINITION OF MEASURES (SIB MATING)

Consider an individual W in the $(t+1)^{\text{th}}$ generation of a sib mating pedigree (Figure 1). Our aim is to determine the inbreeding coefficient F^{t+1}_{ij} of W . This is a probability statement about the uniting gametes $a_Y b_Y, a_Z b_Z$ and is given by $F^{t+1}_{ij} = X_{ij}(a_Y b_Y, a_Z b_Z)$. SCHNELL's (1961) function of inbreeding is $F^{t_{11}}$ and $F^{t_{00}}$ is his panmictic function. The marginal probabilities, $F^{t_{1.}}$ and $F^{t_{.1}}$, are the one locus inbreeding coefficients for the first and second loci, respectively, which satisfy the following equations.

FIGURE 1.—Sib Mating Pedigree.

$$F_{.1}^t = \frac{1}{4} + \frac{1}{2} F_{.1}^{t-1} + \frac{1}{4} F_{.1}^{t-2}$$

$$F_{.1}^t = \frac{1}{4} + \frac{1}{2} F_{.1}^{t-1} + \frac{1}{4} F_{.1}^{t-2} \quad (3)$$

Each of the uniting gametes forming W may be either a parental-type or a recombinant-type gamete formed by an individual (Y or Z) in generation t .

If both gametes are parental, $X_{ij}(a_Y b_Y, a_Z b_Z)$ has one of these four values; $X_{ij}(a_I b_I, a_J b_J)$, $X_{ij}(a_K b_K, a_L b_L)$, $X_{ij}(a_I b_I, a_L b_L)$, $X_{ij}(a_K b_K, a_J b_J)$. The first two involve gametes from one individual and are assigned the symbol θ^t_{ij} while the last two involve gametes from distinct individuals and so are just the inbreeding function F^t_{ij} . For the general measure $X_{ij}(a_r b_t, a_s b_u)$ then, two classes are distinguished:

- F^t_{ij} : $a_r b_t$ and $a_s b_u$ are gametes from distinct individuals in generation $(t-1)$,
 θ^t_{ij} : $a_r b_t$ and $a_s b_u$ are distinct gametes from one individual in generation $(t-1)$.

If one gamete is parental and the other recombinant, $X_{ij}(a_Y b_Y, a_Z b_Z)$ becomes one of these eight measures; $X_{ij}(a_I b_I, a_J b_L)$, $X_{ij}(a_K b_K, a_L b_J)$, $X_{ij}(a_I b_K, a_J b_J)$, $X_{ij}(a_K b_I, a_L b_L)$, $X_{ij}(a_I b_I, a_L b_J)$, $X_{ij}(a_K b_K, a_J b_L)$, $X_{ij}(a_K b_I, a_J b_J)$, $X_{ij}(a_I b_K, a_L b_L)$. The first four cases are similar and are assigned the symbol P^t_{ij} while each of the last four is written Q^t_{ij} . For the general measure $X_{ij}(a_r b_t, a_s b_u)$ then, two further classes are distinguished:

- P^t_{ij} : $a_r b_t$ is a gamete from one individual in generation $(t-1)$, a_s is on the other gamete from that individual, and b_u is on a gamete from the other individual,
 Q^t_{ij} : $a_r b_t$ is a gamete from one individual in generation $(t-1)$, b_u is on the other gamete from that individual, and a_s is on a gamete from the other individual.

When both gametes are recombinant, $X_{ij}(a_Y b_Y, a_Z b_Z)$ takes one of four values; $X_{ij}(a_I b_K, a_J b_L)$, $X_{ij}(a_K b_I, a_L b_J)$, $X_{ij}(a_I b_K, a_L b_J)$, $X_{ij}(a_K b_I, a_J b_L)$. Each of the first two measures are written U^t_{ij} and each of the last two V^t_{ij} . For the general measure $X_{ij}(a_r b_t, a_s b_u)$ then, another two classes are distinguished:

- U^t_{ij} : a_r and a_s are on distinct gametes from one individual and b_t and b_u are on distinct gametes from the other individual in generation $(t-1)$.
 V^t_{ij} : a_r and b_u are on distinct gametes from one individual and a_s and b_t are on distinct gametes from the other individual in generation $(t-1)$.

The six classes of the general measure that have now been defined are those needed in the expansion of F^{t+1}_{ij} back to measures of the previous generation. We will find it convenient to write S^t_{ij} as the average of P^t_{ij} and Q^t_{ij} , and necessary to define one further class, T^t_{ij} , pertaining to such measures as $X_{ij}(a_I b_I, a_K b_L)$:

TABLE 1

Probabilities for the 36 combinations of pairs of a's with pairs of b's from four gametes in any generation of sib mating

	$b_I b_J$	$b_I b_K$	$b_I b_L$	$b_J b_K$	$b_J b_L$	$b_K b_L$
$a_I a_J$	θ	P	P	P	P	U
$a_I a_K$	Q	F	T	T	V	Q
$a_I a_L$	Q	T	F	V	T	Q
$a_J a_K$	Q	T	V	F	T	Q
$a_J a_L$	Q	V	T	T	F	Q
$a_K a_L$	U	P	P	P	P	θ

T_{ij}^t : $a_i b_t$ is a gamete from one individual and a_s and b_u are on distinct gametes from the other individual in generation $(t-1)$.

For the four gametes in any generation there are six pairings of a's in combination with six pairings of b's. The probabilities corresponding to the types of functions are given in Table 1 for the 36 combinations.

The set of measures $[F_{ij}, \theta_{ij}, S_{ij}, T_{ij}, U_{ij}, V_{ij}]^t$ is now complete in the sense that the value of any of them in generation t can be expressed in terms of the set, and only the set, in generation $(t-1)$.

MARGINAL TOTALS

The four components of each of the measures may be arranged in two by two tables as shown in Table 2. Consider P_{ij}^t for example. $P_{1.}^t$ is the value for terms like Prob $[a_I \equiv a_J]$. Now a_I, a_J are the same gene half the time and identical with probability one, and are different genes half the time and identical with probability $F_{1.}^{t-1}$. Thus $P_{1.}^t = (1 + F_{1.}^{t-1})/2$. However, $P_{1.}^t$ is the value for terms like Prob $[b_K \equiv b_J]$ and thus has the value $F_{1.}^t$. When we combine P_{ij}^t and Q_{ij}^t we see that

$$\begin{aligned}
 S_{1.}^t &= \frac{(P_{1.}^t + Q_{1.}^t)}{2} \\
 &= \frac{(\frac{1}{2} + \frac{1}{2} F_{1.}^{t-1} + F_{1.}^t)}{2} \\
 &= F_{1.}^{t+1}.
 \end{aligned} \tag{4}$$

Other values are found similarly.

TRANSITION EQUATIONS

Transition equations are found by straightforward, but sometimes tedious, probability arguments. Using $F_{ij}^{t+1} = X_{ij}(a_Y b_X, a_Z b_Z)$ as an example, we know that parental and recombinant gametes are produced by an individual with probabilities $(1 + \lambda)/2$ and $(1 - \lambda)/2$, respectively, where λ is the linkage

TABLE 2

Marginal totals of probability measures

F_{11}^t F_{01}^t	F_{10}^t F_{00}^t	$F_{1.}^t$ $F_{0.}^t$	θ_{11}^t θ_{01}^t	θ_{10}^t θ_{00}^t	$\frac{1}{2} + \frac{1}{2} F_{1.}^{t-1}$ $\frac{1}{2} F_{0.}^{t-1}$
$F_{.1}^t$	$F_{.0}^t$	1	$\frac{1}{2} + \frac{1}{2} F_{.1}^{t-1}$	$\frac{1}{2} F_{.0}^{t-1}$	1
P_{11}^t P_{01}^t	P_{10}^t P_{00}^t	$\frac{1}{2} + \frac{1}{2} F_{1.}^{t-1}$ $\frac{1}{2} F_{0.}^{t-1}$	Q_{11}^t Q_{01}^t	Q_{10}^t Q_{00}^t	$F_{1.}^t$ $F_{0.}^t$
$F_{.1}^t$	$F_{.0}^t$	1	$\frac{1}{2} + \frac{1}{2} F_{.1}^{t-1}$	$\frac{1}{2} F_{.0}^{t-1}$	1
S_{11}^t S_{01}^t	S_{10}^t S_{00}^t	$F_{1.}^{t+1}$ $F_{0.}^{t+1}$	T_{11}^t T_{01}^t	T_{10}^t T_{00}^t	$F_{1.}^t$ $F_{0.}^t$
$F_{.1}^{t+1}$	$F_{.0}^{t+1}$	1	$F_{.1}^t$	$F_{.0}^t$	1
U_{11}^t U_{01}^t	U_{10}^t U_{00}^t	$\frac{1}{2} + \frac{1}{2} F_{1.}^{t-1}$ $\frac{1}{2} F_{0.}^{t-1}$	V_{11}^t V_{01}^t	V_{10}^t V_{00}^t	$F_{1.}^t$ $F_{0.}^t$
$\frac{1}{2} + \frac{1}{2} F_{.1}^{t-1}$	$\frac{1}{2} F_{.0}^{t-1}$	1	$F_{.1}^t$	$F_{.0}^t$	1

parameter ($\lambda = 0$ for free recombination, and $\lambda = 1$ for zero recombination). Hence the three classes of uniting gametes and their probabilities are

$$\begin{array}{ll}
 \text{parental} \times \text{parental} & \frac{(1 - \lambda)^2}{4} \\
 \text{parental} \times \text{recombinant} & \frac{1 - \lambda^2}{2} \\
 \text{recombinant} \times \text{recombinant} & \frac{(1 - \lambda)^2}{4}
 \end{array} \quad (5)$$

For the $p \times p$ class, the uniting gametes ($a_Y b_Y, a_Z b_Z$) can take with equal probability one of the four possible values ($a_I b_I, a_L b_L$), ($a_J b_J, a_K b_K$), ($a_I b_I, a_J b_J$), ($a_K b_K, a_L b_L$) for which the probabilities of the first two are F_{ij}^t and of the latter two are θ_{ij}^t , with an average of $(F_{ij}^t + \theta_{ij}^t)/2$. The $p \times r$ class of uniting gametes takes with equal probabilities one of eight possible values and the probabilities for these values are divided equally between P_{ij}^t and Q_{ij}^t to give an average of $(P_{ij}^t + Q_{ij}^t)/2 = S_{ij}^t$. Finally for the $r \times r$ class the equally probable four values have an average probability of $(U_{ij}^t + V_{ij}^t)/2$. Multiplying together the probabilities of the classes and the probabilities given the classes leads to the first of the transition equations (6).

The other transition equations are obtained by similar arguments, but the diagram (Figure 1) must be extended to include gametes to the other individual in order to identify the types of functions in generation $t + 1$. The simplest argument is for θ_{ij} , gametes from the same individual, and is the same for any mating

system. It is only the equation for S_{ij}^t that necessitates the use of the probability measure T_{ij}^{t-1} .

The equations (6) are shown in their simplest possible form by making use of the expressions for the marginal totals as shown in Table 2, and are true for i and j taking values 0 or 1. Each equation thus represents four equations. Because of these marginal totals, and the fact that $F_{1.}^t, F_{.1}^t$ satisfy equations (1), it is necessary to solve equations (6) for only one component of each function in order to have solved the whole system. For simplicity we work with the X_{00}^t 's.

$$\begin{aligned}
 F_{ij}^t &= \frac{(1+\lambda)^2}{8} F_{ij}^{t-1} + \frac{(1+\lambda)^2}{8} \theta_{ij}^{t-1} + \frac{1-\lambda^2}{2} S_{ij}^{t-1} + \frac{(1-\lambda)^2}{8} U_{ij}^{t-1} + \frac{(1-\lambda)^2}{8} V_{ij}^{t-1} \\
 \theta_{ij}^t &= ij F_{11}^{t-1} + \frac{i}{2} F_{10}^{t-1} + \frac{j}{2} F_{01}^{t-1} + \frac{1+(-1)^{i+j}\lambda^2}{4} F_{00}^{t-1} \\
 S_{ij}^t &= \frac{1}{4} T_{ij}^{t-1} + ij (S_{11}^{t-1} - \frac{1}{4}) + \frac{i}{2} S_{10}^{t-1} + \frac{j}{2} S_{01}^{t-1} + \frac{1}{4} S_{00}^{t-1} \\
 T_{ij}^t &= \frac{1+\lambda}{8} F_{ij}^{t-1} + \frac{1+\lambda}{8} \theta_{ij}^{t-1} + \frac{1}{2} S_{ij}^{t-1} + \frac{1-\lambda}{8} U_{ij}^{t-1} + \frac{1-\lambda}{8} V_{ij}^{t-1} \\
 U_{ij}^t &= ij V_{11}^{t-1} + \frac{i}{2} V_{10}^{t-1} + \frac{j}{2} V_{01}^{t-1} + \frac{1}{4} V_{00}^{t-1} \\
 V_{ij}^t &= \frac{1}{8} F_{ij}^{t-1} + \frac{1}{8} \theta_{ij}^{t-1} + \frac{1}{2} S_{ij}^{t-1} + \frac{1}{8} U_{ij}^{t-1} + \frac{1}{8} V_{ij}^{t-1}
 \end{aligned} \tag{6}$$

($i, j = 0, 1$)

The six equations for the X_{00}^t 's may be written in matrix form as in equation (7), $\mathbf{u}_t = \mathbf{A}\mathbf{u}_{t-1}$ where $\mathbf{A}$ is the transition matrix and $\mathbf{u}$ is a vector of six elements.

$$\begin{bmatrix} F_{00} \\ \theta_{00} \\ S_{00} \\ T_{00} \\ U_{00} \\ V_{00} \end{bmatrix}_t = \begin{bmatrix} \frac{(1+\lambda)^2}{8} & \frac{(1+\lambda)^2}{8} & \frac{1-\lambda^2}{2} & 0 & \frac{(1-\lambda)^2}{8} & \frac{(1-\lambda)^2}{8} \\ \frac{1+\lambda^2}{4} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ \frac{1+\lambda}{8} & \frac{1+\lambda}{8} & \frac{1}{2} & 0 & \frac{1-\lambda}{8} & \frac{1-\lambda}{8} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{4} \\ \frac{1}{8} & \frac{1}{8} & \frac{1}{2} & 0 & \frac{1}{8} & \frac{1}{8} \end{bmatrix} \begin{bmatrix} F_{00} \\ \theta_{00} \\ S_{00} \\ T_{00} \\ U_{00} \\ V_{00} \end{bmatrix}_{t-1} \tag{7}$$

If the initial members are a pair of outbred, unrelated individuals, the initial vector $\mathbf{u}_1$ is put equal to $[1, \frac{1+\lambda^2}{4}, \frac{1}{2}, 1, \frac{1}{4}, 1]$.

We now wish to use (7) to establish a recurrence formula for F_{00}^t , i.e., we want to express F_{00}^t in terms of only F_{00} 's in previous generations. The simplest such equation is, in general, found by way of the minimal polynomial $f(x)$ of the matrix A . This is the simplest polynomial with $f(A)$ equal to the zero matrix. Hence $f(A) \mathbf{u}_t = \mathbf{0}$, and we have a recurrence formula for the whole vector $\mathbf{u}_t$, and so certainly for one component (F_{00}) of the vector. This ignores the possibility of factorization of $f(x)$ but, in general, for our matrix A the minimal polynomial has no factors and is equal to the characteristic polynomial and thus of degree six. The recurrence formula for F_{00}^t must, therefore, go back six generations, and is now displayed together with the initial six values assuming a non-inbred initial population:

$$F_{00}^{t+6} = \frac{(4+2\lambda+\lambda^2)}{8} F_{00}^{t+5} + \frac{(4-2\lambda+\lambda^2+2\lambda^3+\lambda^4)}{32} F_{00}^{t+4} \\ - \frac{(2+8\lambda+6\lambda^2+6\lambda^3+\lambda^4)}{128} F_{00}^{t+3} - \frac{\lambda(1+2\lambda+2\lambda^3+\lambda^4)}{256} F_{00}^{t+2} \\ + \frac{\lambda(1+5\lambda^2+2\lambda^4)}{1024} F_{00}^{t+1} - \frac{\lambda^3(1+\lambda^2)}{2048} F_{00}^t. \quad (8)$$

$$F_{00}^0 = 1$$

$$F_{00}^1 = 1$$

$$F_{00}^2 = \frac{18+3\lambda^2+2\lambda^3+\lambda^4}{32}$$

$$F_{00}^3 = (100+16\lambda^2+22\lambda^3+17\lambda^4+4\lambda^5+\lambda^6)/256 \quad (9)$$

$$F_{00}^4 = (512+112\lambda^2+124\lambda^3+138\lambda^4+76\lambda^5+45\lambda^6+14\lambda^7+3\lambda^8)/2048$$

$$F_{00}^5 = (2074+704\lambda^2+800\lambda^3+848\lambda^4+592\lambda^5+532\lambda^6+306\lambda^7+133\lambda^8+32\lambda^9+5\lambda^{10})/16384.$$

Using F_{00}^2 and $F_{.1}^2 = F_{.1}^2 = \frac{1}{4}$ leads to $F_{11}^2 = (2+3\lambda^2+2\lambda^3+\lambda^4)/32$ as given by previous authors (HALDANE (1937), SCHNELL (1961)).

In two special cases, equation (8) reduces considerably. For complete linkage ($\lambda=1$) the characteristic polynomial of the matrix factorizes to:

$$(x - \frac{1}{2})(x - \frac{1}{4})(x + \frac{1}{4})(x + \frac{1}{8})(x^2 - \frac{1}{2}x - \frac{1}{4}),$$

while for zero linkage ($\lambda=0$) it factorizes to:

$$x^3(x + \frac{1}{4})(x^2 - \frac{3}{4}x + 1/16).$$

In such cases, we take the smallest number of factors (so that the recurrence formula goes back the least number of generations) to form the polynomial $g(x)$ such that $g(A)\mathbf{u}_t$ has zero first component. By doing so for complete linkage we get equation (10) which is satisfied by $F_{0.}^t$ and $F_{.0}^t$, as we expect for two genes being transmitted as one unit. For zero linkage we get equation (11) which is satisfied by $(F_{0.}^t, F_{.0}^t)$, which we also expect since now the two genes are transmitted independently.

$$F_{00}^t = \frac{1}{2} F_{00}^{t-1} + \frac{1}{4} F_{00}^{t-2}. \quad (10)$$

$$F_{00}^t = \frac{1}{2} F_{00}^{t-1} + \frac{1}{8} F_{00}^{t-2} - 1/64 F_{00}^{t-3}. \quad (11)$$

Of course these two results may be derived without having to go to the sixth degree characteristic equation. For complete linkage, the complete set is just $[F_{ij}, \theta_{ij}]^t$ and θ_{ij} may be eliminated easily for all values of i and j . For zero linkage we see that $F_{ij}^t = T_{ij}^t = V_{ij}^t$ and $\theta_{ij}^t = U_{ij}^t$ so that the complete set is $[F_{ij}, S_{ij}, U_{ij}]^t$. Here again it is simple, especially for $i = j = 0$, to eliminate T_{ij}^t and U_{ij}^t .

As t increases, the ratio between successive values of F_{00}^t tends to the largest characteristic root of matrix A . When $\lambda = 1$ this value is $\varepsilon = (1 + \sqrt{5})/4$ as in FISHER (1949), and when $\lambda = 0$ the value is ε^2 . Intermediate values are given in Table 3, where values marked with an asterisk have also been found by BENNETT (1954).

Numerical values of F_{11}^t are given in Table 4. Calculations were made for F_{00}^t and the one locus inbreeding coefficients assuming that there was no initial inbreeding at either locus, and F_{11}^t was computed as follows:

$$F_{11}^t = F_{00}^t + F_{1.}^t + F_{.1}^t - 1. \quad (12)$$

APPROXIMATION

An approximation can be made which simplifies the recurrence formula considerably, yet yields good numerical results. We replace $P_{ij}^t; Q_{ij}^t$ with the values they would have for unlinked loci and replace V_{ij}^t by F_{ij}^t . Now we are able to write

TABLE 3

Largest characteristic root of transition matrix

λ	Root
0.00	0.65451*
0.05	0.65459
0.10	0.65484
0.15	0.65530
0.20	0.65600*
0.25	0.65699
0.30	0.65832
0.35	0.66007
0.40	0.66231
0.45	0.66517
0.50	0.66879
0.55	0.67334
0.60	0.67906*
0.65	0.68625
0.70	0.69524*
0.75	0.70645
0.80	0.72031*
0.85	0.73724
0.90	0.75758*
0.95	0.78150
1.00	0.80902*

TABLE 4

Numerical values of F_{11}^t for selected values of λ

t	0.0	0.1	0.2	0.3	0.4	λ 0.5	0.6	0.7	0.8	0.9	1.0
2	.06250	.06350	.06680	.07288	.08324	.09570	.11380	.13738	.16730	.20450	.25000
3	.14053	.14134	.14392	.14915	.15800	.17169	.19169	.21976	.25797	.30877	.37500
4	.25000	.25061	.25279	.25721	.26483	.27701	.29563	.32321	.36320	.42014	.50000
5	.35254	.35302	.35475	.35826	.36440	.37444	.39027	.41476	.45216	.50878	.59375
6	.45142	.45179	.45311	.45580	.46055	.46843	.48118	.50165	.53449	.58721	.67188
7	.53931	.53959	.54058	.54261	.54621	.55225	.56224	.57880	.60654	.65363	.73438
8	.61647	.61668	.61741	.61891	.62159	.62614	.63379	.64684	.66964	.71046	.78516
9	.68256	.68271	.68325	.68434	.68631	.68969	.69546	.70558	.72395	.75859	.82617
10	.73853	.73863	.73902	.73982	.74125	.74372	.74803	.75577	.77053	.79927	.85938
15	.90490	.90492	.90499	.90513	.90539	.90587	.90676	.90854	.91254	.92263	.95126
20	.96650	.96651	.96652	.96654	.96658	.96666	.96683	.96720	.96816	.97124	.98311
30	.99595	.99595	.99595	.99595	.99595	.99595	.99596	.99597	.99602	.99626	.99797
40	.99951	.99951	.99951	.99951	.99951	.99951	.99951	.99951	.99952	.99953	.99976
50	.99994	.99994	.99994	.99994	.99994	.99994	.99994	.99994	.99994	.99994	.99997

$$\begin{aligned}
 P_{00}^t &= \text{Prob} [a_I \neq a_J, b_I \neq b_L] \\
 &\cong \text{Prob} [a_I \neq a_J] \text{Prob} [b_I \neq b_L] \\
 &= \frac{1}{2} F_{0,0}^{t-1} F_{0,0}^t
 \end{aligned}
 \tag{13}$$

$$\begin{aligned}
 Q_{00}^t &= \text{Prob} [a_I \neq a_L, b_I \neq b_J] \\
 &\cong \text{Prob} [a_I \neq a_L] \text{Prob} [b_I \neq b_J] \\
 &= \frac{1}{2} F_{0,0}^t F_{0,0}^{t-1} \\
 V_{00}^t &\cong F_{00}^t,
 \end{aligned}$$

which, from the first of equation (6) leads to:

$$\begin{aligned}
 F_{00}^t &= \frac{1+\lambda^2}{4} F_{00}^{t-1} + \frac{2+3\lambda^2+2\lambda^3+\lambda^4}{32} F_{00}^{t-2} \\
 &\quad + \frac{1-\lambda^2}{8} (F_{0,0}^{t-1} F_{0,0}^{t-2} + F_{0,0}^{t-2} F_{0,0}^{t-1}).
 \end{aligned}
 \tag{14}$$

While this approximate equation holds exactly for zero or complete linkage, it is also good for intermediate linkage as can be seen by comparing the results in Table 5 with those in Table 4.

DISCUSSION

Inbreeding at one locus can be elucidated completely for any form of inbreeding by digametic probability measures, F and θ , F for uniting gametes and θ for random gametes from one or two individuals. To treat inbreeding at two linked loci requires the introduction of trigametic and quadrigametic probability measures when the system of mating involves four or more gametes from one generation to the next. Selfing, which has only two transitional gametes, can be accom-

TABLE 5
Approximate values of F_{11}

t	0.1	0.3	λ 0.5	0.7	0.9
2	.06350	.07288	.09570	.13738	.20450
3	.14141	.14961	.17249	.22043	.30892
4	.25063	.25729	.27693	.32256	.41925
5	.35298	.35783	.37307	.41212	.50635
6	.45172	.45513	.46633	.49763	.58342
7	.53951	.54184	.54981	.57393	.64874
8	.61661	.61817	.62372	.64174	.70491
9	.68265	.68370	.68749	.70068	.75277
10	.73859	.73928	.74185	.75133	.79349
15	.90491	.90499	.90533	.90688	.91917
20	.96651	.96652	.96656	.96678	.96983
30	.99595	.99595	.99595	.99595	.99611
40	.99951	.99951	.99951	.99951	.99952
50	.99994	.99994	.99994	.99994	.99994

modated for any number of loci with just digametic probability measures. For sib mating it was found that two types of trigametic, S and T , and two types of quadrigametic, U and V , probability measures were sufficient, in conjunction with the digametic ones, to provide transitional equations from one generation to the next. Not all types of tri- and quadrigametic probability measures necessary for a general pedigree analysis of two locus inbreeding are encountered in sib mating, however. These are to be developed elsewhere.

Some generalizations are clear at this point. The x -gametic probability measures required for a joint treatment of n linked loci are $x = 2, 3, \dots, 2n$, but in the analysis of systems of mating x is limited such that $x = 2, 3, \dots, 2N$, where N is the number of individuals. Thus, for selfing and any number of loci, $x = 2$; for sib mating, $x = 2$ for $n = 1$ and $x = 2, 3, 4$ for $n \geq 2$; for double first cousins, $x = 2$ for $n = 1$, $x = 2, 3, 4$ for $n = 2$, $x = 2, 3, 4, 5, 6$ for $n = 3$ and $x = 2, 3, 4, 5, 6, 7, 8$ for $n \geq 4$; and so on. What is not clear is the number of types required for measures of each order x , as related to the system of mating. It is the total number of measures which determines the rank of the transitional matrix, e.g., A in (7). While we do not at this point see how to generalize the rank of A for sib mating and any number of loci, the number of measures for a particular number of loci can be found by the same method demonstrated for two loci.

The set of measures in (7) is the fewest for transition from one generation to the next. They may be reduced to smaller sets. For example, one may eliminate θ , T , and U to provide a set of three measures but the transition formulae go back three generations. In essence it is the reduction of the set into one measure which leads to the recurrence formula (6) but which goes back six generations.

The application of the recurrence formulae (6), (7) and (8) is not limited to the initial conditions (9) used in the numerical example. Genes may be spe-

cified to be identical at either or both loci in either or both initial individuals, and in various ways between individuals. Initial measures must be adjusted accordingly, and different values of the same type of measure are averaged. This points out the fact that averages have been used in two different ways. For example, consider the four F probabilities in Table 1. These are identical to each other regardless of the initial conditions. On the other hand, the two θ 's may be initially different, and if so the arithmetic average is to be used. The same applies to the other measures when they are initially different. If all four genes at one locus of the two initial individuals are identical by descent then all X_{00}^1 's are zero, and inbreeding at the other locus is accounted for by the equations in (6).

It should be stressed that linkage affects only joint probabilities, F_{11} , F_{10} , F_{01} and F_{00} , of identity by descent of genes, and in no way affects the inbreeding coefficient, F_1 , or $F_{.1}$. One must reflect then on the purposes to be served by the joint probabilities, of which two come readily to mind. One has to do with the distribution of multiple zygotes and the other has to do with the variance of "identity by descent" among linked loci (SCHNELL 1961). Complete information for the variance is provided by two locus probabilities.

To see the effect on the distribution of multiple zygotes, let

$$\begin{aligned}\delta_{11}^t &= F_{11}^t - F_{.1}^t F_{.1}^t \\ \delta_{10}^t &= F_{10}^t - F_{.1}^t F_{.0}^t \\ \delta_{01}^t &= F_{01}^t - F_{.0}^t F_{.1}^t \\ \delta_{00}^t &= F_{00}^t - F_{.0}^t F_{.0}^t\end{aligned}\quad (15)$$

and note that

$$\delta_{11}^t = \delta_{00}^t = -\delta_{10}^t = -\delta_{01}^t. \quad (16)$$

In Table 4 the values under $\lambda = 0.0$ are equivalent to $F_{.1}^t F_{.1}^t$. By comparing corresponding values for other λ 's we see that $\delta_{11} \geq 0$, that it increases with λ , that it is maximum for any λ in one of the first three generations and that it decreases in time to zero. Thus the effect of linkage is to increase the frequency of double homozygotes over that expected for genes which recombine freely. At the same time, using (16), we note that the frequency of double heterozygotes is affected the same way and that the frequencies of the single heterozygotes are reduced correspondingly. The fact that all δ 's go to zero with time is a reflection of the approach to fixation at both loci.

Similar effects of linkage, but more pronounced, are found with selfing (NARAIN 1965) where

$$\begin{aligned}\delta_{00}^t &= F_{00}^t - F_{.0}^t F_{.0}^t \\ &= \left(\frac{1+\lambda^2}{4}\right)^t - \left(\frac{1}{4}\right)^t,\end{aligned}\quad (17)$$

$\delta_{00} = \delta_{11}$ being maximum in the first generation.

Identity disequilibrium, i.e., δ 's $\neq 0$, is a general result for linkage and any form of inbreeding. It affects but does not specify linkage disequilibrium which depends, in addition, on initial linkage disequilibrium. One can easily deduce for populations which are initially in linkage equilibrium that identity disequilibrium

leads to a correlation between genotypes of the two loci, but that the effect is transient and disappears as the individuals become homozygous.

SUMMARY

Sib mating with two linked loci is studied. Trigametic and quadrigametic probability measures of identity by descent are introduced in addition to the usual digametic ones. A minimal set of six measures is required for transition from one generation to the next. When reduced to the panmictic function the recurrence relation goes back six generations. The panmictic function in conjunction with the inbreeding coefficient provides the inbreeding function. Numerical values of the inbreeding function, of the largest characteristic root of the transition matrix, and of an approximation to the inbreeding function are given. —Some generalizations concerning the effects of the system of mating and number of loci on the types of probability measures required, concerning the effects of linkage on the inbreeding function and concerning the utility of the inbreeding function are discussed.

LITERATURE CITED

- BENNETT, J. H., 1954 The distribution of heterogeneity upon inbreeding. *J. Roy. Statist. Soc. B* **16**: 88–99.
- FISHER, R. A., 1949 *Theory of Inbreeding*. Oliver and Boyd Ltd., Edinburgh.
- HALDANE, J. B. S., 1937 Some theoretical results of continued brother-sister mating. *J. Genet.* **34**: 265–274. — 1949 The association of characters as a result of inbreeding and linkage. *Ann. Eugenics* **15**: 15–23.
- HALDANE, J. B. S., and C. H. WADDINGTON, 1931 Inbreeding and linkage. *Genetics* **16**: 357–374.
- MALÉCOT, G., 1948 *Les Mathématiques de L'Hérédité*. Masson et Cie., Paris.
- NARAIN, P., 1965 Homozygosity in a selfed population with an arbitrary number of linked loci. *J. Genet.* **59**: 254–266.
- SCHNELL, F. W., 1961 Some general formulations of linkage effects in inbreeding. *Genetics* **46**: 947–957.
- SHIKATA, M., 1962 The generalized inbreeding coefficient in a Markov process. *Rep. Statist. Appl. Res.*, JUSE **9**: 127–136. — 1965a A generalization of the inbreeding coefficient. *Biometrics* **21**: 665–681. — 1965b Limit of application of generalized inbreeding coefficient and generalized coefficient of parentage. *Ann. Repts. Research Group Biophys. Japan* **5**: 65–66.
- WRIGHT, S., 1921 Systems of mating II. *Genetics* **6**: 124–143.