Estimation of Recombination Frequencies and Construction of RFLP Linkage Maps in Plants From Crosses Between Heterozygous Parents

E. Ritter, C. Gebhardt and F. Salamini

Max-Planck-Institut für Züchtungsforschung, D-5000 Köln 30, West Germany

Manuscript received January 5, 1990
Accepted for publication March 26, 1990

ABSTRACT

The construction of a restriction fragment length polymorphism (RFLP) linkage map is based on the estimation of recombination frequencies between genetic loci and on the determination of the linear order of loci in linkage groups. RFLP loci can be identified as segregations of singular or allelic DNA-restriction fragments. From crosses between heterozygous individuals several allele (fragment) configurations are possible, and this leads to a set of formulas for the evaluation of p, the recombination frequency between two loci. Tables and figures are presented illustrating a general outline of gene mapping using heterozygous populations. The method encompasses as special cases the mapping of loci from segregating populations of pure lines. Formulas for deriving the recombination frequencies and information functions are given for different fragment configurations. Information functions derived for relevant configurations are also compared. A procedure for map construction is presented, as it has been applied to RFLP mapping in an allogamous crop.

WITH the discovery of a new marker class termed "restriction fragment length polymorphisms" (RFLP), marker based selection is currently receiving attention and support in crop breeding (reviewed in BECKMANN and SOLLER 1986). RFLP linkage maps have been constructed for several crop species including maize, tomato, lettuce, rice and potato (HELENTJARIS *et al.* 1986; BERNATZKY and TANKSLEY 1986; HELENTJARIS 1987; LANDRY *et al.* 1987; ZAMIR and TANKSLEY 1988; MCCOUCH *et al.* 1988; BONIERBALE, PLAISTED and TANKSLEY 1988; GEBHARDT *et al.* 1989). Moreover, RFLP markers are virtually unlimited in numbers, the only restriction to the efficiency of this technique being the DNA sequence divergence between the genotypes tested.

Restriction fragments of nuclear DNA varying in length between parental genotypes are detected by Southern blot hybridization to cloned homologous sequences as probes (SOUTHERN 1975). Any single polymorphic restriction fragment segregates as a codominant Mendelian marker in the progeny from parents being heterozygous for that fragment. The distance on the linkage map between any two RFLP markers is determined by measuring the recombination frequency. Linked markers are aggregated in linkage groups. The linear order of markers within each linkage group is deduced from the genetic distances relative to each other in two-, three- or multiple-point estimates. The number of linkage groups is equivalent to the chromosome number of the species. Most RFLP maps in plants have been obtained from segregating populations, F_2 and/or backcrosses, derived from homozygous inbred lines (*e.g.*, HELENTJARIS *et al.* 1986; BERNATZKY and TANKSLEY 1986).

We have recently produced a RFLP map for the potato (*Solanum tuberosum* ssp. *tuberosum*) (GEBHARDT *et al.* 1989). In the diploid state, potato clones are self-incompatible and characterized by a high genetic load. Both conditions preclude the possibility of obtaining pure lines. In the case of our map, two highly heterozygous parents were crossed to obtain a segregating offspring. In the paper we describe the theoretical background for RFLP linkage analysis from any type of F_1 populations, including those from heterozygous individuals, which encompasses as special cases mapping in F_2 and backcross populations from homozygous inbred lines.

METHODS AND RESULTS

Calculation of recombination frequencies between loci defined by single restriction fragments: This situation has been considered as separated from the case of loci defined by the existence of allelic restriction fragments (see later). In the case of a locus defined only by a single fragment A, care is not taken to individuate possible fragments allelic to the same locus: the presence of A is scored versus its absence in a progeny segregating for A. Allelic states to A are therefore scored as null (O = no alternative fragment). Genotypes having the same phenotype (A present) may be homozygous (AA) or heterozygous (AO), where A behaves as a dominant marker. In a F1 cross of the type AO × OO, the segregation ratio is 1:1 (presence vs. absence), while the F_1 of a cross AO ×

TABLE 1

Derivation of recombination frequencies

A. Single Fragment Loci

Fragment configuration of parents		Mating table for an AB/OO type (coupling)				
$\frac{AB}{OO} \times \frac{OO}{OO}$	GF	GF	$\frac{1-p}{2}$	$\frac{1-p}{2}$	$\frac{p}{2}$	$\frac{p}{2}$
	GF	GT	*OO*	*OO*	*OO*	*OO*
	$\frac{1-p'}{2}$	*AB*	AB	AB	AB	AB
	$\frac{1-p'}{2}$	*OO*	OO	OO	OO	OO
	$\frac{p'}{2}$	*AO*	AO	AO	AO	AO
	$\frac{p'}{2}$	*OB*	OB	OB	OB	OB

	Distribution of phenotypes	AB : AO : OB : OO
	Absolute linkage	8 : 0 : 0 : 8
	Absence of linkage	4 : 4 : 4 : 4

$p = p'$ Calculation table

Phenotypes	p_j	$\frac{\delta p_j}{\delta p}$	$\frac{1}{p_j}\frac{\delta p_j}{\delta p}$	$\frac{1}{p_j}\left(\frac{\delta p_j}{\delta p}\right)^2$	Z_j
AB	$\frac{1-p}{2}$	$\frac{-1}{2}$	$\frac{-1}{1-p}$	$\frac{1}{2(1-p)}$	*a*
AO	$\frac{p}{2}$	$\frac{1}{2}$	$\frac{1}{p}$	$\frac{1}{2p}$	*b*
OB	$\frac{p}{2}$	$\frac{1}{2}$	$\frac{1}{p}$	$\frac{1}{2p}$	*c*
OO	$\frac{1-p}{2}$	$\frac{-1}{2}$	$\frac{-1}{1-p}$	$\frac{1}{2(1-p)}$	*d*
Sum	1	0		$i = \frac{1}{p(1-p)}$	*n*

Maximum likelihood equation

$$\frac{-a}{1-p} + \frac{b}{p} + \frac{c}{p} + \frac{-d}{1-p} = 0 \quad \rightarrow P = \frac{b+c}{n}$$

AO will segregate 3:1. Segregation at a second RFLP locus *B* can be defined accordingly.

Presence and absence of fragments A and B can be arranged in 2^8 configurations in the four loci available for two diploid parents, including cases of homo- besides those of heterozygosity. The best estimate *P* for the recombination frequency *p* between A and B can be obtained by use of the maximum likelihood method of FISHER (1921) and requires a specific treatment for each of the parental fragment configurations. As an example, the derivation of *P* for the configuration AB/OO×OO/OO (fragments A and B are both heterozygous and present on the same chromosome only in one parent) is shown in Table 1A.

The mating table shows the gamete types (GT), their expected frequencies (GF) as functions of *p* and the phenotypes of the progeny resulting from crossing the parent AB/OO with OO/OO. In case of absolute linkage between A and B ($p = 0$) the two parental phenotypes are expected in the F_1 progeny with a frequency of 50% each, whereas in the absence of linkage ($p = 0.5$) four phenotypes (two parental, two recombinant) with 25% frequency each appear. Recombination frequencies between A and B can only be estimated based on the difference between the expected phenotypic frequencies for absolute linkage and for an independent segregation of the two markers.

The Chi square test (MATHER 1938) will establish whether the observed numbers of phenotypes (Z_j) deviate significantly from those expected in case of independent segregation. If A and B are supposed to be linked, the recombination frequency *p* can be estimated by solving the maximum likelihood Equation 1

B. Loci with Allelic Fragments

Fragment configuration of parents		Mating table for $\alpha\beta/\alpha\beta$ type (coupling)[a]				
$\frac{A_1B_1}{A_2B_2} \times \frac{A_1B_1}{A_2B_2}$		GF	$\frac{1-p}{2}$	$\frac{1-p}{2}$	$\frac{p}{2}$	$\frac{p}{2}$
	GF	GT	A_1B_1	A_2B_2	A_1B_2	A_2B_1
	$\frac{1-p'}{2}$	A_1B_1	A_1B_1	$A_1A_2B_1B_2$	$A_1B_1B_2$	$A_1A_2B_1$
	$\frac{1-p}{2}$	A_2B_2	$A_1A_2B_1B_2$	A_2B_2	$A_1A_2B_2$	$A_2B_1B_2$
	$\frac{p'}{2}$	A_1B_2	$A_1B_1B_2$	$A_1A_2B_2$	A_1B_2	$A_1A_2B_1B_2$
	$\frac{p'}{2}$	A_2B_1	$A_1A_2B_1$	$A_2B_1B_2$	$A_1A_2B_1B_2$	A_2B_1

	Distribution of phenotypes	A_1B_1	A_1B_2	$A_1B_1B_2$	A_2B_1	A_2B_2	$A_2B_1B_2$	$A_1A_2B_1$	$A_1A_2B_2$	$A_1A_2B_1B_2$
	Absolute linkage	4	0	0	0	4	0	0	0	8
	Absence of linkage	1	1	2	1	1	2	2	2	4

$p = p'$ — Calculation table

Phenotypes	p_j	$\frac{\delta p_j}{\delta p}$	$\frac{1}{p_j}\frac{\delta p_j}{\delta p}$	$\frac{1}{p_j}\left(\frac{\delta p_j}{\delta p}\right)^2$	$-Z_j$
A_1B_1	$\frac{(1-p)^2}{4}$	$\frac{p-1}{2}$	$\frac{-2}{1-p}$	1	$a1$
A_1B_2	$\frac{p^2}{4}$	$\frac{p}{2}$	$\frac{2}{p}$	1	$a2$
$A_1B_1B_2$	$\frac{p(1-p)}{2}$	$\frac{1-2p}{2}$	$\frac{1-2p}{p(1-p)}$	$\frac{(1-2p)^2}{2p(1-p)}$	$a3$
A_2B_1	$\frac{p^2}{4}$	$\frac{p}{2}$	$\frac{2}{p}$	1	$a4$
A_2B_2	$\frac{(1-p)^2}{4}$	$\frac{p-1}{2}$	$\frac{-2}{1-p}$	1	$a5$
$A_2B_1B_2$	$\frac{p(1-p)}{2}$	$\frac{1-2p}{2}$	$\frac{1-2p}{p(1-p)}$	$\frac{(1-2p)^2}{2p(1-p)}$	$a6$
$A_1A_2B_1$	$\frac{p(1-p)}{2}$	$\frac{1-2p}{2}$	$\frac{1-2p}{p(1-p)}$	$\frac{(1-2p)^2}{2p(1-p)}$	$a7$
$A_1A_2B_2$	$\frac{p(1-p)}{2}$	$\frac{1-2p}{2}$	$\frac{1-2p}{p(1-p)}$	$\frac{(1-2p)^2}{2p(1-p)}$	$a8$
$A_1A_2B_1B_2$	$\frac{1-2p+2p^2}{2}$	$-1(1-2p)$	$\frac{-2(1-2p)}{1-2p+2p^2}$	$\frac{2(1-2p)^2}{1-2p+2p^2}$	$a9$
Sum	1	0		$i = \frac{2(1-3p+3p^2)}{p(1-p)(1-2p+2p^2)}$	n

Maximum likelihood equation

$$\frac{-2(a1+a5)}{1-p} + \frac{(1-2p)(a3+a6+a7+a8)}{p(1-p)} + \frac{2(a2+a4)}{p} + \frac{-2(1-2p)a9}{1-2p+2p^2} = 0$$

GF = gametic frequency; GT = gamete type; p,p' = recombination frequency of male and female gametes; Z_j = observed numbers of phenotypes.

[a] See text.

(FISHER 1921)

$$\frac{\delta \ln L(p)}{\delta p} = \sum_j Z_j \frac{1}{p_j} \frac{\delta p_j}{\delta p} = 0 \qquad (1)$$

where p_j are the expected frequencies and Z_j the observed numbers of phenotypes. Here and in the following Equation 2, the terms needed for the solution are calculated as examplified in the calculation table (Table 1A). The information fuction I_p which measures the quality of the estimate P is given by (MATHER 1938)

$$-\mathrm{E}\left(\frac{\delta^2 \ln L(p)}{\delta p^2}\right) = n \sum_j \frac{1}{p_j}\left(\frac{\delta p_j}{\delta p}\right)^2 = ni = I_p \qquad (2)$$

where n is the sample size (=number of offspring). The variance of P is then given by

$$V(P) = 1/I_p \qquad (3)$$

and the standard deviation by $\sqrt{V(P)}$. The maximum likelihood estimator is a minimum variance unbiased estimator of the recombination frequency p (RAO 1952).

In Table 1A the expected frequencies p_j (first column) are obtained by multiplying the gamete frequencies giving rise to a specific phenotype and summing up the products over the mating table. For example

$$p_{AB} = \left(\frac{1-p}{2}\frac{1-p'}{2}\right) + \left(\frac{1-p}{2}\frac{1-p'}{2}\right) + \left(\frac{p}{2}\frac{1-p'}{2}\right) + \left(\frac{p}{2}\frac{1-p'}{2}\right) = \frac{1-p}{2} \qquad (4)$$

with the male frequency of recombinant gametes (p) equal to that of the female (p'). The other terms (columns 2, 3, and 4) are derivatives of p_j. Using the calculation table, Equation 1 is formulated as

$$\frac{-a}{1-p} + \frac{b}{p} + \frac{c}{p} + \frac{-d}{1-p} = 0$$

and solving for p gives the estimate

$$P = \frac{b+c}{a+b+c+d} = \frac{b+c}{n}$$

with

$$\hat{V}(P) = \frac{P(1-P)}{n}$$

from Equations 2 and 3.

If in a cross only four phenotypes are present, as it is with single fragment loci, it may be convenient (see below) to estimate p with the product formula of FISHER and BALMAKUND (1928) which is easy to calculate (IMMER 1930). Thus p is estimated by $\check{P}$ solving the equation:

$$\frac{p_{AB}\, p_{OO}}{p_{AO}\, p_{OB}} = \frac{Z_{AB}\, Z_{OO}}{Z_{AO}\, Z_{OB}} \qquad (5)$$

with p_j as expected frequencies and Z_j as observed numbers of phenotypes. If the variance is the same as with the maximum likelihood method then the product formula gives a fully efficient estimate of p (BAILEY 1961).

Similar as shown in Table 1A, mating tables can be assembled for all the 2^8 possible fragment configurations at the loci A and B of two diploid parents. In crosses these fragment configurations originate a maximum of four phenotypes because the homozygous or heterozygous states for a fragment cannot be distinguished. However, out of the 256 configurations, only a few have expected phenotypic frequencies differing between absolutely linked and unlinked fragments A and B, and these are therefore useful for linkage analysis. They are combined in three types:

1. The AB/OO-type with the configurations AB/OO × OO/OO (coupling, see Table 1A) and AO/OB × OO/OO (repulsion), characterized by the presence of both fragments A and B in one parent and absence in the other;

2. The AB/AO with the configurations AB/OO × AO/OO (coupling) and AO/OB × AO/OO (repulsion) in which one fragment is present in both parents and the other only in one;

3. The AB/AB type with the configurations AB/OO × AB/OO (coupling), AO/OB × AO/OB (repulsion) and AB/OO × AO/OB (coupling/repulsion) with both fragments shared by the parents.

For each informative fragment configuration as defined above, a calculation table can be developed by expressing the expected phenotypic frequency p_j as a function of p, obtained from the mating table as exemplified in Equation 4, and by calculating the partial derivatives and the other terms necessary for solving the maximum likelihood Equation 1. In doing this, three assumptions are made:

1. The recombination frequency during gamete formation is the same in both parents ($p = p'$);

2. Reciprocal crosses result in the same phenotypic frequencies (P1 × P2 = P2 × P1);

3. The phenotypic frequencies are identical and independent of which homologous chromosomes are paired (*e.g.*, AB/OO = OO/AB).

Table 2A summarizes the formulas necessary to calculate the recombination frequency estimators P and $\check{P}$ for the seven usable fragment configurations of two single fragment loci. The formulas for the AB/OO and AB/AB type were derived by solving equation (1), while for the AB/AO type equation (5) was used due to its lower computational complexity. The solution of the maximum likelihood equation:

$$\frac{-a}{2-p} + \frac{b}{1+p} + \frac{c}{p} + \frac{-d}{1-p} = 0$$

(AB/AO coupling)

would in fact lead in this case to a polynomial of third order, while the application of the product formula gives a quadratic equation, the variance being the same in both cases. Analogous results can be obtained for repulsion. With the product formula the value of X, as defined in Table 2A, is always larger or equal to one. The formula is not defined for $X = 1$ ($ad = bc$) or the denominator being equal to zero [$bc = 0$ (coupling), $ad = 0$ (repulsion)]. If, however, X approaches one or the denominator approaches zero, then the estimate of p converges upon 0.5 and zero respectively. Similar conclusions can be drawn for the other cases, when the product formula is applied.

DISTORTED SEGREGATION RATIOS

In the F_1 a deviation from the segregation ratio of 1:1 for fragments contributed only by one parent and from 3:1 for fragments present in both parents may result due to a reduced viability of some of the resulting phenotypes (reduced viability of certain gametes is not considered here). Significant deviations from the normal ratios are detected with the Chi square test (summarized in MATHER 1938). If the "skewing factor" u (ratio of the phenotypes with and without a fragment A) is considered, the phenotypic frequencies in the mating table of Table 1A can be expressed as

$$p_{AB} = u(1 - p)/(u + 1),\ p_{OO} = (1 - p)/(u + 1),\ p_{AO} = up/(u + 1) \text{ and } p_{OB} = p/(u + 1)$$

summing to 1 (BAILEY 1961). When only one fragment shows distorted segregation, u disappears in subsequent calculations and the estimate for p is the same as with segregating fragments without distortion. Nevertheless the variance must be specifically calculated because it is different from the case of absence of distortion. If both fragments are distorted and the "skewing factor" for B is v then the phenotypic frequencies are expressed as

$$p_{AB} = uv(1 - p)/D;\ p_{OO} = (1 - p)/D;$$

$$p_{AO} = up/D \text{ and } p_{OB} = vp/D$$

with

$$D = uv(1 - p) + p(u + v) + 1 - p.$$

The estimate for p results in complex maximum likelihood equations, but using the product formula as suggested by BAILEY (1961), solutions can be found for the AB/OO type and the AB/AB type (see $\check{P}*$ and $\check{S}*$ in Table 2A). For the AB/AO type, the estimation formula for p is always the same whether distorted segregation ratios are observed or not, since the product formula is used in all cases.

CALCULATION OF RECOMBINATION FREQUENCIES BETWEEN LOCI DEFINED BY ALLELIC RESTRICTION FRAGMENTS

If two fragments A_1 and A_2 are detected with the same probe and if they are linked 100% in repulsion ($p = 0$), they can be treated as allelic fragments (although they do not have to be so in the molecular sense). In a progeny of heterozygous parents, a locus may therefore be represented by up to four codominant allelic fragments in the combinations A_1A_3, A_1A_4, A_2A_3, A_2A_4 if A_1 and A_2 are the alleles of P1 and A_3 and A_4 of P2. If only two alleles are present in both parents (both parents have, for instance, A_1A_2) and knowing these alleles based on their electrophoretic pattern, the homozygous or heterozygous state of a locus can be deduced. Recombination frequencies between two such loci are derived by the procedure described for single fragment loci. As an example, the fragment configuration $A_1B_1/A_2B_2 \times A_1B_1/A_2B_2$, with A_1, A_2 and B_1, B_2 being allelic fragments of two loci A and B, is shown in Table 1B. As seen in the mating table, nine phenotypes can be distinguished and their frequencies vary according to the linkage intensity between the loci A and B. The terms necessary to formulate Equations 1 and 2 are given in the calculation table. The maximum likelihood equation is a polynomial of higher order, that can be solved iteratively using, for example, Newton's approximation method. Table 2B summarizes the formulas and maximum likelihood equations (where a universal solution is not possible) for the estimation of p between two loci with allelic fragments (Nos. 1–3) or for mixed situations where linkages between a single fragment locus and a locus with allelic fragments are considered (Nos. 4–6). Allelic configurations at loci with allelic fragments are here indicated introducing the additional letters α and β ($\alpha = A_1/A_2$, $\alpha' = A_1/A_3$, $\beta = B_1/B_2$, $\beta' = B_1/B_3$), and their configurations in a cross are defined in terms of allelic states in Table 3. The configurations $\alpha\beta$/OO, αB/OO (Nos. 1a and 4b) and $\alpha\beta/\alpha$O, αB/αO, αB/BO (Nos. 1b, 4a and 5) have similar solutions for P as AB/OO ad AB/AO respectively (Table 2A). For the configurations $\alpha\beta/\alpha\beta$, $\alpha\beta/\alpha'\beta'$ and αB/αB, respectively (Nos. 2, 3 and 6) the three cases of coupling, repulsion, and coupling/repulsion have been considered.

In the three allelic configuration $\alpha\beta/\alpha'\beta'$ (No. 3) the sixteen genotypes can be distinguished, allowing a very precise estimate of p (see also Figure 1). A configuration where four different fragments are found at a locus is treated as the $\alpha\beta/\alpha'\beta'$ configuration by attaching corresponding genotypes.

In a similar way as described in Table 1 further mating and calculation tables could be set up considering three and more loci, where several parameters have to be estimated.

INFORMATIVITY OF P DEPENDS ON THE FRAGMENT CONFIGURATION

The information function I_p or its reciprocal, the variance $V(P)$ (Equations 2 and 3), is a measure of the precision of the estimated recombination frequency P (MATHER 1938). Table 4 lists the information func-

TABLE 2

Recombination frequencies and maximum likelihood equations for different fragment configurations

A. Single fragment loci

Type	Recombination frequency
	$P = \frac{Z_1 + Z_2}{n}$
AB/OO	$\check{P}^{*\cdot **} = \frac{\sqrt{Z_1 * Z_2}}{\sqrt{Z_1 * Z_2} + \sqrt{Z_3 * Z_4}}$
Coupling	$Z_1 = b\ Z_2 = c\ Z_3 = a\ Z_4 = d$
Repulsion	$Z_1 = a\ Z_2 = d\ Z_3 = b\ Z_4 = c$
AB/AO	$\check{P}^{**} = \frac{x+3}{2(1-x)} + \sqrt{\left(\frac{x+3}{2(1-x)}\right)^2 - \frac{2}{1-x}}$
Coupling	with $x = \frac{ad}{bc}$
Repulsion	with $x = \frac{bc}{ad}$
AB/AB	
Coupling	$P = 1 - \sqrt{S}$
Repulsion	$P = \sqrt{S}$
Coupling/repulsion	$P = \frac{1}{2} - \sqrt{\frac{1-4S}{4}}$
with	$S = X + \sqrt{X^2 + \frac{2d}{n}}$
and with	$X = \frac{a - 2b - 2c - d}{2n}$
or with	$\check{S}^{*\cdot **} = \frac{X+1}{X-1} - \sqrt{\frac{3X+1}{(X-1)^2}}$
and with	$X = \frac{ad}{bc}$ for coupling
or	$X = \frac{bc}{ad}$ for repulsion, coupling/repulsion

tions for different fragment configurations, which comprise classical allelic configurations from backcross (AB/OO type) and F_2 (AB/AB, $\alpha\beta/\alpha\beta$, $\alpha\beta/\alpha'\beta'$ types) derived from inbred lines, and which have been analyzed by MATHER (1938) and ALLARD (1956). All fragment configurations of the AB/AO type do not occur in a F_2 or backcross and are here reported as original contributions.

Figure 1, homologous to the ones of MATHER (1938) and ALLARD (1956), compares graphically the information functions and shows the relative values of I_p as dependent from the recombination frequency p. In the figure, functions from 1 to 6 of Table 4 are represented as quotients, with the I_p function of the AB/OO type (No. 1) being the common divisor. The relative I_p function of the AB/OO type has therefore a constant value of 1 in the figure. These relative functions are now independent from the number of offspring n. An I_p value of 0.5 means that with the same number of progeny (n) the estimated recombination frequency P for a specific fragment configuration is only half as precise as it would be for the AB/OO configuration. Recombination frequencies calculated for loci with allelic fragments are more precise than those obtained for single fragment loci. In this respect the best configuration is $\alpha\beta/\alpha'\beta'$ (three alleles at both loci), being always twice as good as the AB/OO type. The $\alpha\beta/\alpha\beta$ type is nearly as good as the $\alpha\beta/\alpha'\beta'$ type at low recombination frequencies but its precision decreases with increasing p. Whereas the AB/AB coupling type is similarly informative as AB/OO at low recombination frequencies, the informativity of the AB/AB repulsion and the AB/AO type is considerably lower ranging from O to 44% and 25% to 33%, respectively, compared to the AB/OO type.

FORMATION OF RFLP LINKAGE GROUPS

By estimating the recombination frequency between any two polymorphic restriction fragments, RFLP loci can be ordered in groups of linkage. Within each group the linear order is deduced from all cal-

B. Loci with allelic fragments

No.	Type	Recombination frequency		
1				
a)	$\alpha\beta/OO$ $P = \frac{Z_1 + Z_2}{n'}$	Coupling	$Z_1 = A_1B_2$	$Z_2 = A_2B_1$
b)	$\alpha\beta/\alpha O$ with $n' = A_1B_1 + A_2B_2 + A_1B_2 + A_2B_1$	Repulsion	$Z_1 = A_1B_1$	$Z_2 = A_2B_2$
2	$\alpha\beta/\alpha\beta$ Maximum likelihood equation $\frac{-2(Z_1 + Z_5)}{1-p} + \frac{(1-2p)(Z_3 + Z_6 + Z_7 + Z_8)}{p(1-p)} + \frac{2(Z_2 + Z_4)}{p} + \frac{-2(1-2p)Z_9}{1-2p+2p^2} = 0$			
a)	Coupling $Z_i = a_i$ (see Table 1B)			
b)	Repulsion $Z_1 = a_4$ $Z_5 = a_2$ (other $Z_i = a_i$)			
c)	Coupling/repulsion $P = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{(a_1 + a_2 + a_4 + a_5 + a_9)}{2n}}$			
3	$\alpha\beta/\alpha'\beta'$ $P = \frac{2X_2 + X_3 + X_4}{2n}$			
a)	Coupling $X_1 = A_1B_1 + A_1A_3B_1B_3 + A_1A_2B_1B_2 + A_2A_3B_2B_3$ $X_2 = A_1B_2B_3 + A_1A_3B_1B_2 + A_1A_2B_1B_3 + A_2A_3B_1$ $X_3 = A_1B_1B_3 + A_1A_3B_1 + A_1A_2B_2B_3 + A_2A_3B_1B_2$ $X_4 = A_1B_1B_2 + A_1A_3B_2B_3 + A_1A_2B_1 + A_2A_3B_1B_3$			
b)	Repulsion $X_1 = X_2$, $X_2 = X_1$			
c)	Coupling/repulsion (2 subtypes, depending on mating table)			
4				
a)	$\alpha B/\alpha O$ $P = \frac{Z_1 + Z_2}{n'}$	Coupling	$Z_1 = A_1O$	$Z_2 = A_2B$
b)	$\alpha B/OO$ with $n' = A_1B_1 + A_2B_2 + A_1B_2 + A_2B_1$	Repulsion	$Z_1 = A_2O$	$Z_2 = A_1B$
5	$\alpha\beta/BO$			
a)	Coupling	as AB/A-Type ($a \Rightarrow A_1B$, $b \Rightarrow A_2B$, $c \Rightarrow A_1O$, $d \Rightarrow A_2O$)		
b)	Repulsion			
6	$\alpha B/\alpha B$ Maximum likelihood equation $\frac{-2pZ_1}{1-p^2} + \frac{2Z_2}{p} + \frac{2(1-p)Z_3}{p(2-p)} + \frac{-2Z_4}{1-p} + \frac{-(1-2p)Z_5}{1-p+p^2} + \frac{(1-2p)Z_6}{p(1-p)} = 0$			
a)	Coupling $Z_1 = A_1B$, $Z_2 = A_1O$, $Z_3 = A_2B$, $Z_4 = A_2O$, $Z_5 = A_1A_2B$, $Z_6 = A_1A_2O$			
b)	Repulsion $Z_1 = A_2B$, $Z_2 = A_2O$, $Z_3 = A_1B$, $Z_4 = A_1O$, $Z_5 = A_1A_2B$, $Z_6 = A_1A_2O$			
c)	Coupling/repulsion Maximum likelihood equation $\frac{Z_1 + Z_3}{1-p+p^2} - \frac{Z_2 + Z_4}{p(1-p)} - \frac{2Z_5}{1+2p-2p^2} + \frac{2Z_6}{1-2p+2p^2} = 0$ $Z_1 = A_1B$, $Z_2 = A_1O$, $Z_3 = A_2B$, $Z_4 = A_2O$, $Z_5 = A_1A_2B$, $Z_6 = A_1A_2O$			

* Calculation for both fragments with distorted segregations. ** calculated with the product formula (see text for details). $a = Z_{AB}$, $b = Z_{AO}$, $c = Z_{OB}$, $d = Z_{OO}$ (as defined in Table 1). n = number of offspring. See text and Table 3 for fragment configurations of parents.

culated recombination frequencies among the loci considered. In Table 5 the order A–B–C–D of four fragments is supported by the given relationships expressed by the P values. If a two point estimate order of linked fragments is determined irrespective of the fragment configuration, contradictions may occur because the precision and the limits of significance of the estimate of p depend on the information function used. Ambiguities may also arise for the configuration AB/AB, because cases of repulsion and of coupling/

TABLE 3

Definitions of configurations with single restriction fragments and allelic fragments

No.	Configuration name		Configuration
	AB/OO	Coupling	AB/OO × OO/OO
		Repulsion	AO/OB × OO/OO
	AB/AO	Coupling	AB/OO × AO/OO
		Repulsion	AO/OB × AO/OO
	AB/AB	Coupling	AB/OO × AB/OO
		Repulsion	AO/OB × AO/OB
		Coupling/repulsion	AB/OO × AO/OB
1 a)	αβ/OO	Coupling	A_1B_1/A_2B_2 × OO/OO
		Repulsion	A_1B_2/A_2B_1 × OO/OO
1 b)	αβ/αO	Coupling	A_1B_1/A_2B_2 × A_1O/A_2O
		Repulsion	A_1B_2/A_2B_1 × A_1O/A_2O
2 a)	αβ/αβ	Coupling	A_1B_1/A_2B_2 × A_1B_1/A_2B_2
b)		Repulsion	A_1B_2/A_2B_1 × A_1B_2/A_2B_1
c)		Coupling/repulsion	A_1B_1/A_2B_2 × A_1B_2/A_2B_1
3 a)	αβ/α′β′	Coupling	A_1B_1/A_2B_2 × A_1B_1/A_3B_3
b)		Repulsion	A_1B_2/A_2B_1 × A_1B_3/A_3B_1
c)		Coupling/repulsion	A_1B_1/A_2B_2 × A_1B_3/A_3B_1
			A_1B_2/A_2B_1 × A_1B_1/A_3B_3
4 a)	αB/αO	Coupling	A_1B/A_2O × A_1O/A_2O
		Repulsion	A_1O/A_2B × A_1O/A_2O
b)	αB/OO	Coupling	A_1B/A_2O × OO/OO
		Repulsion	A_1O/A_2B × OO/OO
5	αB/BO	Coupling	A_1B/A_2O × OB/OO
		Repulsion	A_1O/A_2B × OB/OO
6 a)	αB/αB	Coupling	A_1B/A_2O × A_1B/A_2O
b)		Repulsion	A_1O/A_2B × A_1O/A_2B
c)		Coupling/repulsion	A_1B/A_2O × A_1O/A_2B

repulsion cannot be distinguished based on the distribution of phenotypes. For example, with a distribution of phenotypes where AB = 51%, AO = 24%, OB = 24% and OO = 1%, a recombination frequency of 20% is calculated in case of adoption of the AB/AB configuration in repulsion, but only of 4% in the case of the AB/AB coupling/repulsion. Difficulties can be overcome by first setting up linkage subgroups based on fragment configurations taking into account only single fragment loci. The subgroups can then be combined in a linkage group by using available fragment positions as due to the existence of loci with allelic fragments. This is illustrated in Figure 2 for a hypothetical linkage group. The homologous chromosomes of the parents P1 and P2 characterized by segregating restriction fragments at 14 RFLP loci with single and/or allelic fragments are reported in Figure 2A. Using the configurations AB/OO (coupling and repulsion) and AB/AB (coupling only) which have the highest informativity for single fragment loci, five linkage subgroups are constructed (Figure 2B). The gene order within these subgroups can now be determined by multiple-point estimates using available computer programmes like "MAPMAKER" (LANDER *et al.* 1987). However MAPMAKER cannot be applied to a whole dataset, where fragments segregating with 3:1 and 1:1 ratios are intermixed.

TABLE 4

Information functions

No.	Type of configuration	Information function I_p
1	AB/OO	$I_p = \frac{n}{p(1-p)}$
2	AB/AO	$I_p = \frac{n(1+2p-2p^2)}{2p(1-p^2)(2-p)}$
3	AB/AB coupling	$I_p = \frac{2n(3-4p+2p^2)}{p(2-p)(3-2p+p^2)}$
4	AB/AB repulsion	$I_p = \frac{2n(1+2p^2)}{(2+p^2)(1-p^2)}$
5	αβ/α′β′ coupling repulsion	$I_p = \frac{2n}{p(1-p)}$
6	αβ/αβ coupling repulsion	$I_p = \frac{2n(1-3p+2p^2)}{p(1-p)(1-2p+2p^2)}$

n = Number of offspring. p = Recombination frequency.

The subgroups are combined into a linkage group with 'allelic bridges' by searching for a genetic locus present in two different subgroups and with alleles segregating in both parents. The allelic state of the fragments considered is detected using configurations of the types AB/AO or AB/AB (repulsion) and searching for absence of recombination ($p = 0$). These configurations have lower informativity than, for instance, AB/OO and AB/AB (coupling) which however, have been used to establish the primary subgroups. In this way, one subgroup established by using AB/OO configurations can be combined with a second subgroup by an allelic fragment of the locus common to both parents and forming an AB/AO repulsion type of configuration (see locus *L12* in Figure 2). Two AB/AB coupling subgroups are matched by an "allelic bridge" of the AB/AB repulsion configuration (locus *L9*) and even two AB/OO subgroups of linked loci each specific for the two parents are combined with an AB/AB coupling subgroup by an allelic bridge of two AB/AO repulsion types (locus *L4*). The relative orientation of the subgroups is deduced from comparing recombination frequencies of AB/AO or AB/AB configurations (repulsion) of consecutive loci. Figure 2C shows the combined linkage group with 11 single fragment loci (five from P1, two from P2 and four from both parents) and three loci with allelic fragments.

DISCUSSION

With the analysis of RFLP patterns in segregating populations two types of fragments can appear: dominant single restriction fragments and codominant allelic restriction fragments. Depending on the degree of heterozygosity of the parents, a certain percentage

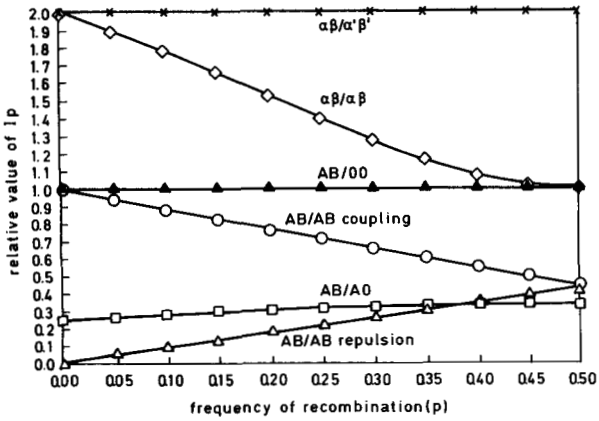

FIGURE 1.—Information functions of six different fragment configurations relative to the I_p-function of the AB/OO configuration and their dependence from the recombination frequency p. The functions are independent from the number of offspring (see text). The formulas of Table 4 have been used.

TABLE 5

Relationships between P values for four loci with the order A–B–C–D

	A	B	C
B	P_{AB}	—	
	∧		
C	$P_{AC} > P_{BC}$		—
	∧	∧	
D	$P_{AD} > P_{BD} > P_{CD}$		

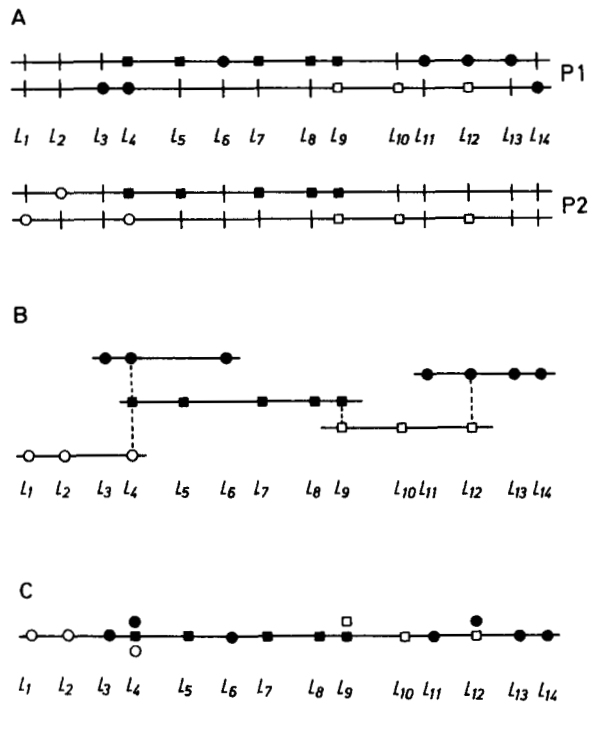

FIGURE 2.—An example of linkage map construction in a progeny from heterozygous parents. A, Hypothetical fragment distribution on two homologous chromosomes of two parents P1 and P2. B, Linkage subgroups and their connection via "allelic bridges." C, Final linkage group. (●) AB/OO linkage subgroup from P1. (○) AB/OO linkage subgroup from P2. (■, □) AB/AB (coupling) linkage subgroups. L_i, locus name.

of polymorphic markers cannot be analyzed either due to homozygosity of both parents of an F_1, or due to loss of heterozygous polymorphic fragment in the specific F_1 plant used to generate a backcross population.

All different types of fragment configurations and the corresponding formulas and ML-equations for calculating the recombination frequencies between two loci, which can occur in RFLP analysis, are summarized in Table 2. These fragment configurations encompass as special cases classical allelic configurations from backcross and F_2 populations derived from inbred lines. Fragment configurations specifically introduced for RFLP analysis in this paper include the AB/AO, $\alpha\beta$/OO, $\alpha\beta/\alpha$O, αB/αO, αB/OO and αB/BO types, as well as the coupling/repulsion subtypes of the AB/AB configuration and Nos. 2, 3 and 6 of Table 2B. As evident from Table 2, identical formulas can apply to different fragment configurations.

As developed in this paper, map construction in allogamous plant species for which only heterozygous individuals are available can also make use of fragment configurations based on single restriction fragments. In particular, the AB/AO and the AB/AB (repulsion) configurations are of interest because they are useful for the identification of allelic bridges and for selecting the correct orientations among independently derived linkage subgroups. If for a polymorphism allelic restriction fragments are present, more precise estimates of p can be obtained (see Figure 1). Thus the mapping strategy illustrated in Figure 2, as supported from the formulas given in Tables 1 and 2, provides a general instrument of RFLP linkage analysis. In its computerized form this strategy made possible the mapping of polymorphic fragments on the 12 chromosomes of potato (GEBHARDT *et al.* 1989).

The data base were x segregating restriction fragments with m probes scored for presence or absence in the parents and n offspring. Recombination frequencies were estimated depending on the fragment configuration for each pair of fragments taking into account distorted segregation ratios. As shown in Figure 1, however, the precision of the estimate of p between single fragment loci is lower compared with composite loci of the $\alpha\beta/\alpha\beta$ type used for linkage analysis in F_2 populations (HELENTJARIS *et al.* 1986). An advantage of the method, however, is the possibility of including in the analysis distorted segregation ratios without introducing new parameters (see Table 2A).

LITERATURE CITED

ALLARD, R. W., 1956 Formulas and tables to facilitate the calculation of recombination values in heredity. Hilgardia **24:** 235–278.

BAILEY, N. T. J., 1961 *Introduction to the Mathematical Theory of Linkage.* Clarendon Press, Oxford.

BECKMANN, J. S., and M. SOLLER, 1986 Restriction fragment length polymorphisms in plant genetic improvement. Oxf. Surv. Plant Mol. Cell Biol. **3:** 197–250.

BERNATZKY, R., and S. D. TANKSLEY, 1986 Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics **112:** 887–898.

BONIERBALE, M. W., R. L. PLAISTED and S. D. TANKSLEY, 1988 RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics **120:** 1095–1103.

FISHER, R. A., 1921 On the mathematical foundations of theoretical statistics. Philos. Trans. R. Soc Lond A **122:** 309–368.

FISHER, R. A., and B. BALMAKUND, 1928 The estimation of linkage from the offspring of selfed heterozygotes. J. Genet. **20:** 79–92.

GEBHARDT, C., E. RITTER, T. DEBENER, U. SCHACHTSCHABEL, B. WALKEMEIER, H. UHRIG and F. SALAMINI, 1989 RFLP-analysis and linkage mapping in *Solanum tuberosum.* Theor. Appl. Genet. **78:** 65–75.

HELENTJARIS, T., 1987 A genetic linkage map for maize based on RFLPs. Trends Genet. **3:** 217–221.

HELENTJARIS, T., M. SLOCUM, S. WRIGHT, A. SCHAEFER and J. NIENHUIS, 1986 Construction of genetic linkage maps in maize and tomato using restriction fragment length polymorphisms. Theor. Appl. Genet. **72:** 761–769.

IMMER, F. R., 1930 Formulae and tables for calculating linkage intensities. Genetics **15:** 81–98.

LANDER, E. S., P. GREEN, J. ABRAHAMSON, A. BARLOW, M. J. DALY, S. E. LINCOLN and L. NEWBURG, 1987 Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

LANDRY, B. S., R. V. KESSELI, B. FARRARA and R. W. MICHELMORE, 1987 A genetic linkage map of lettuce (*Lactuca sativa* L.) with restriction fragment length polymorphism, isozyme, disease resistance and morphological markers. Genetics **116:** 331–337.

MCCOUCH, S. R., G. KOCHERT, Z. H. YU, Z. Y. WANG, G. S. KHUSH, W. R. COFFMAN and S. D. TANKSLEY, 1988 Molecular mapping of rice chromosomes. Theor. Appl. Genet. **76:** 815–829.

MATHER, K., 1938 *The Measurement of Linkage in Heredity.* Methuen & Co., London.

RAO, R. C., 1952 *Advanced Statistical Methods in Biometric Research.* John Wiley & Sons, New York.

SOUTHERN, E. M., 1975 Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:** 503–517.

ZAMIR, D., and S. D. TANKSLEY, 1988 Tomato genome is comprised largely of fast-evolving, low copy-number sequences. Mol. Gen. Genet. **213:** 254–261.

Communicating editor: S. D. TANKSLEY