

THE LINKAGE PROBLEM IN A TETRASOMIC WILD PLANT, *LYTHRUM SALICARIA*

BY *R. A. FISHER*

DEPARTMENT OF GENETICS, UNIVERSITY OF CAMBRIDGE

LINKAGE in polysomic organisms is more complex than that in disomic organisms, broadly for three reasons.

(1) The multiplicity of the segregating genotypes. These may be enumerated and represented by the partitions of the number four for tetrasomics or six for hexasomics, in two or more dimensions, according to the number of linked loci involved. Thus, with two linked loci in a tetrasomic organism, there have been shown (FISHER, 1947) to be 24 sets of isomorphic genotypes, of which 8 are self-conjugate, or unaltered by interchange of loci, and the remainder constitute 8 pairs of conjugate sets.

If only two different genes are available at each locus, as in the linkage of short versus non-short styles and purple versus rosy in *Lythrum salicaria*, we may make up 19 doubly heterogenic genotypes belonging to four self-conjugate sets, and one pair of conjugates. Of these the most important for the present work are the set of 4, including bisimplex coupling, RS/rs/rs/rs, and the second set of 4, including bisimplex repulsion, Rs/rS/rs/rs. (Table 1.)

(2) The second element of complexity in tetrasomic linkage is that instead of there being for each pair of loci only two modes of gamete formation, namely the recombinant and the non-recombinant gametes, there are with two loci in a tetrasomic organism 11 modes of gamete formation possible, the relative frequencies of which are not given by theory, but which it is a matter of experimental enquiry to determine.

For l linked loci with tetrasomics the number of distinguishable modes of gamete formation is not 2^{l-1} as with disomic organisms, but may be enumerated by the expression

$$\frac{1}{48}16^l + \frac{1}{3}4^l + \frac{1}{3},$$

giving the numbers 2, 11, 107, 1451 for cases with 1, 2, 3 and 4 loci respectively. (Table 2.)

TABLE 1. *Doubly heterogenic genotypes.*

Bipartition

	R r		
s	— 3	Set of 4 genotypes, such as	RS/rs/rs/rs Bisimplex coupling
S	1 —		
	R r		
s	1 2	Set of 4 genotypes	Rs/rS/rs/rs Bisimplex repulsion
S	— 1		
	R r	R r	
s	1 2	— 2	Conjugate sets of 4 each
S	1 —		
			$\left\{ \begin{array}{l} \text{RS/Rs/rs/rs} \\ \text{RS/rS/rs/rs} \end{array} \right.$ Simplo-duplex
	R r		
s	— 2	Set of 2 genotypes	$\left\{ \begin{array}{l} \text{RS/RS/rs/rs} \\ \text{Rs/Rs/rS/rS} \end{array} \right.$ Biduplex coupling and Repulsion
S	2 —		
	R r		
s	1 1	Set of 1 genotype	RS/Rs/rS/rs Biduplex neutral
S	1 1		

TABLE 2. *Number of modes of gamete formation with l linked loci.*

Disomic	$2^l - 1$			
Tetrasomic	$\frac{1}{4}16^l + \frac{1}{3}4^l + \frac{1}{8}$			
	1	2	3	4
	$\frac{1}{4}16^l$	$\frac{1}{3}$	$5\frac{1}{3}$	$85\frac{1}{3}$
	$\frac{1}{3}4^l$	$1\frac{1}{3}$	$5\frac{1}{3}$	$21\frac{1}{3}$
		$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
	$\frac{1}{4}16^l + \frac{1}{3}4^l + \frac{1}{8}$	<u>2</u>	<u>11</u>	<u>107</u>
			<u>107</u>	<u>1451</u>
x_1	1	2	6	22
$3x_1^2 - 1$	<u>2</u>	<u>11</u>	<u>107</u>	<u>1451</u>

$$x_1 = 4x_1 - 1 - 2$$

The eleven modes of formation for two loci are distinguished and classified in Table 3.

(3) In the third place the frequencies of the different gametic genotypes in the gametic series of any given plant cannot be determined

TABLE 3. *Eleven modes of gamete formation for two linked loci.*

Genotype $a_1b_1/a_2b_2/a_3b_3/a_4b_4$.					
Type of Gamete	Number	Formed by bivalents	Recombinant Chromosomes	Double reduction at	
				A	B
1 a_1b_1/a_2b_2	6	+	0	.	.
2 a_1b_1/a_1b_1	4	.	0	+	+
3 a_1b_1/a_2b_3	24	+	$\frac{1}{2}$.	.
4 a_1b_1/a_1b_2	12	.	$\frac{1}{2}$	+	.
5 a_1b_1/a_2b_1	12	.	$\frac{1}{2}$.	+
6 a_1b_2/a_3b_4	12	+	1	.	.
7 a_2b_1/a_3b_1	12	.	1	.	+
8 a_1b_2/a_1b_3	12	.	1	+	.
9 a_1b_2/a_2b_3	24	.	1	.	.
10 a_1b_2/a_1b_2	12	.	1	+	+
11 a_1b_2/a_2b_1	6	.	1	.	.

directly by backcrossing to a multiple recessive, since doubly dominant offspring, *i. e.* short purples, will be produced by any of the 5 gametic genotypes

- RS/RS
- RS/Rs
- RS/rS
- RS/rs
- Rs/rS.

Even if dominance were completely absent the last two of these, having the same genic content, would be indistinguishable in the first generation.

The gametic series of any chosen plant can, however, be obtained experimentally by using the offspring of the first backcross to perform a second backcross to the double recessive, in which case phenotypically indistinguishable doubly dominant plants may give distinguishable progenies. The present paper reports some results from a number of such second backcross progenies.

An accurate determination of the gametic output of known genotypes does not immediately supply the frequencies of the 11 modes of gamete formation possible. For any chosen genotype, such as bi-simplex coupling, one may set out in accordance with MENDEL's law the frequencies of gametes produced by each of these 11 modes. Generally speaking, however, we may expect any particular gametic genotype to be produced by more than one mode of gamete formation from a given parental genotype. The law of inheritance in such a case

TABLE 4. *Gametic matrix for bisimplex coupling, and similar genotypes.*

Mode of gamete formation	$\frac{AB}{AB}$	$\frac{AB}{Ab}$	$\frac{AB}{aB}$	$\frac{AB}{ab}$	$\frac{Ab}{aB}$	$\frac{Ab}{Ab}$	$\frac{Ab}{ab}$	$\frac{aB}{aB}$	$\frac{aB}{ab}$	$\frac{ab}{ab}$
1. a_1b_1/a_2b_2	.	.	.	6	6
2. a_1b_1/a_1b_1	3	9
3. a_1b_1/a_2b_3	.	.	.	3	.	.	3	.	3	3
4. a_1b_1/a_1b_2	.	3	3	6
5. a_1b_1/a_2b_1	.	.	3	.	.	.	3	.	.	6
6. a_1b_2/a_3b_4	6	.	6	.
7. a_2b_1/a_3b_1	6	3	.	3
8. a_1b_2/a_1b_3	3	.	.	6	3
9. a_1b_2/a_2b_3	3	.	3	.	3	3
10. a_1b_2/a_1b_2	3	.	3	.	6
11. a_1b_2/a_2b_1	6	6

Deficiency matrix

	1	2	3	4	5	6	7	8	9	10	11
1	.	.	-2	.	.	1
2	1	.	.	-2	.	1
3	1	-1	-1	.	1	.

TABLE 5. *Gametic matrix for bisimplex repulsion, and similar genotypes.*

Mode of formation	$\frac{AB}{AB}$	$\frac{AB}{Ab}$	$\frac{Ab}{aB}$	$\frac{AB}{ab}$	$\frac{Ab}{aB}$	$\frac{Ab}{AB}$	$\frac{Ab}{ab}$	$\frac{aB}{aB}$	$\frac{aB}{ab}$	$\frac{ab}{ab}$
1. a_1b_1/a_2b_2	2	.	4	.	4	2
2. a_1b_1/a_1b_1	3	.	3	.	6
3. a_1b_1/a_2b_3	.	.	.	1	2	.	3	.	3	3
4. a_1b_1/a_1b_2	.	1	.	.	.	2	.	.	5	4
5. a_1b_1/a_2b_1	.	.	1	.	.	.	5	2	.	4
6. a_1b_2/a_3b_4	.	.	.	2	2	.	2	.	2	4
7. a_2b_1/a_3b_1	.	.	2	.	.	.	4	1	.	5
8. a_1b_2/a_1b_3	.	2	.	.	.	1	.	.	4	5
9. a_1b_2/a_2b_3	.	.	.	2	1	.	3	.	3	3
10. a_1b_2/a_1b_2	1	2	.	2	.	7
11. a_1b_2/a_2b_1	.	.	.	2	.	.	4	.	4	2

Deficiency matrix

	1	2	3	4	5	6	7	8	9	10	11
1	.	.	-2	.	.	1
2	1	.	.	-2	.	1
3	1	-1	-1	.	.	-1

must be expressed in terms of a gametic matrix, with 11 rows, corresponding with the modes of gamete formation, and 10 columns, for the 10 possible gametic genotypes. The appropriate gametic matrices for bisimplex coupling and bisimplex repulsion are given in the tables. (Tables 4 and 5.) It is clear that known frequencies for the different genotypes of gamete possible could not be sufficient, using only a single type of parent, to ascertain the frequencies of all modes of gamete formation. The number of independent equations available for these 11 unknowns is given by the rank of the matrix, and this is found to be 8 for both of the two matrices required. With each of them, therefore, we associate what is called the deficiency matrix, *i. e.* a matrix of three rows and 11 columns, constructed so as to premultiply the gametic matrix to zero, or, in other words, such that its rows represent arbitrary vectors, which can be added to the frequencies of the 11 modes of gamete formation without affecting the gametic output. They indicate in what respects these frequencies are indeterminate.

It is noticeable that the deficiency matrices of the two genotypes employed have only two rows in common, so that using both of these genotypes in the same experimental enquiry there will remain only two degrees of indeterminateness in the estimation of the modes of gamete formation.

A convenient method of dealing with this indeterminateness is to replace the 11 unknowns,

$$f_1, f_2 \dots f_{11},$$

representing the frequencies of the 11 modes of gamete formation and adding necessarily to zero, by 9 linear functions of these:

$$\begin{array}{c} 2f_1 + f_3 \\ f_2 \\ f_4 \\ f_5 \\ f_6 - f_1 - f_{11} \\ f_7 \\ f_8 \\ f_{10} \\ 2f_{11} + f_9 \end{array}$$

also adding to unity, which the two chosen genotypes, bisimplex coupling and bisimplex repulsion, are competent to estimate. This is

done very simply by deleting the first and last rows of the gametic matrices and ascribing the third row to the compound frequency

$$2f_1 + f_3$$

the sixth row to

$$f_6 - f_1 - f_{11}$$

and the ninth row to

$$2f_{11} + f_9.$$

The frequencies associated with these different rows are then only those capable of estimation.

In our practical procedure we have performed the second back-cross only on doubly dominant, or purple short-styled plants. Progenies have not been grown from the two singly dominant classes, the rosy short-styled and the purple long or mid-styled plants from the first generation progenies. To allow for this incompleteness in the experimental analysis of the gametic output it is only necessary to throw together the columns associated with the pairs of gametic genotypes, namely

$$Rs/Rs, Rs/rs$$

and

$$rS/rS, rS/rs,$$

which have not been distinguished. We then have two matrices, each of eight columns and nine rows, which together afford an estimation of

TABLE 6. *Condensed to show the information available from the experiments as actually conducted.*

Frequencies to be estimated	<i>Bisimplex Coupling</i> 8 types of gamete, experimentally distinguished							
	$\frac{AB}{AB}$	$\frac{AB}{Ab}$	$\frac{AB}{aB}$	$\frac{AB}{ab}$	$\frac{Ab}{aB}$	$\frac{Ab}{.b}$	$\frac{aB}{a.}$	$\frac{ab}{ab}$
$2f_1 + f_3$.	.	.	3	.	3	3	3
f_2	3	9
f_4	.	3	3	6
f_5	.	.	3	.	.	3	.	6
$f_6 - f_1 - f_{11}$	6	6	.
f_7	6	3	3
f_8	3	6	3
f_{10}	3	3	6
$f_9 + 2f_{11}$	3	3	3	3
Total	a	b	c	d	e	f	g	h

TABLE 7. Condensed to show the information available from the experiments as actually conducted.

Bisimplex Repulsion
8 types of gamete, experimentally distinguished

Frequencies to be estimated	$\frac{AB}{AB}$	$\frac{AB}{Ab}$	$\frac{AB}{aB}$	$\frac{AB}{ab}$	$\frac{Ab}{aB}$	$\frac{Ab}{.b}$	$\frac{aB}{a.}$	$\frac{ab}{ab}$
$2f_1 + f_3$.	.	.	1	2	3	3	3
f_2	3	3	6
f_4	.	1	.	.	.	2	5	4
f_5	.	.	1	.	.	5	2	4
$f_6 - f_1 - f_{11}$.	.	.	2	2	2	2	4
f_7	.	.	2	.	.	4	1	5
f_8	.	2	.	.	.	1	4	5
f_{10}	1	2	2	7
$f_9 + 2f_{11}$.	.	.	2	1	3	3	3
Total	a'	b'	c'	d'	e'	f'	g'	h'

the nine frequency compounds arrived at above. These reduced matrices are shown in Tables 6 and 7, the total frequencies with which different gametic genotypes are produced being represented by the letters *a, b, c, d, e*, for the five gametic genotypes containing both dominants, *f, g, h*, for the remainder, with the same letters using primes for the corresponding frequencies for the bisimplex repulsion.

Using only frequencies arising from the second backcross, namely the 10 observed values

$$a \dots \dots e$$

$$a' \dots \dots e'$$

direct estimation is now possible for the 9 frequencies required. (Table 8.)

It is important that these are all based on doubly dominant plants, so that the estimation avoids disturbance through inequality of the ratio of dominants to recessives, such as is commonly observed when large numbers are grown. As was expected, moreover, the sum of the 9 estimated frequencies was appreciably greater than unity, the excess

TABLE 8. Estimation from experimental data, using second backcross frequencies only.

$$2f_1 + f_3 = 4d$$

$$f_2 = 4a$$

$$f_4 = 4b$$

$$f_5 = 4c$$

$$f_9 + 2f_{11} = 4e$$

$$f_6 - f_1 - f_{11} = 3(d' + e' - d - e)$$

$$f_7 = 6c' - 2c$$

$$f_8 = 6b' - 2b$$

$$f_{10} = 12a'$$

TABLE 9. *Frequencies from 148 second backcross progenies.*

	Bisimplex			% ₀	
	coupling	repulsion		Unreduced	Reduced
RS/RS	0	0	$2f_1 + f_3$	136.216	118.800
RS/Rs	1.135	0.874	f_2	0	0
RS/rS	0.567	0.291	f_4	4.540	3.960
RS/rs	34.054	3.206	f_5	2.270	1.980
Rs/RS	1.135	19.819	$f_6 - f_1 - f_{11}$	— 36.492	— 31.826
			f_7	0.512	0.534
			f_8	2.974	2.594
			f_{10}	0	0
			$f_9 + 2f_{11}$	4.540	3.960
			Total	114.660	100.002

TABLE 10. *Comparison of frequencies expected and observed.*

	Coupling		Repulsion	
	Exp.	Obs.	Exp.	Obs.
RS/RS	0	0	0	0
RS/Rs	0.990	1.135	0.757	0.874
RS/rS	0.495	0.567	0.254	0.291
RS/rs	29.700	34.054	5.256	3.206
Rs/RS	0.990	1.135	14.826	19.819
Rs/Rs, Rs/rs	16.187	17.724	27.265	26.658
rS/rS, rS/rs	17.197	15.398	28.275	28.792
rs/rs	34.442	29.986	23.365	20.360
	100.001	99.999	99.998	100.000

TABLE 11. *Calculation of recombination and double reduction.*

	Estimate	Factor	Product	Double Reduction at	
				R	S
$2f_1 + f_3$	118.800	$\frac{1}{2}$	59.400	—	—
f_2	0	0	0	0	0
f_4	3.960	$\frac{1}{2}$	1.980	3.960	—
f_5	1.980	$\frac{1}{2}$	0.990	—	1.980
$f_6 - f_1 - f_{11}$	— 31.826	1	— 31.826	—	—
f_7	0.534	1	0.534	—	0.534
f_8	2.594	1	2.594	2.594	—
f_{10}	0	1	0	0	0
$f_9 + 2f_{11}$	3.960	1	3.960	—	—
			37.632	6.554	2.514

being a reflection of the excess of double dominants in the first backcross progenies. The estimates are therefore divided by this experimentally observable factor. The numerical values obtained are shown in

Table 9, and Table 10 exhibits the comparison of the frequencies observed and expected from the two parental genotypes employed, using for both the same frequencies of gamete formation.

Table 11 shows the use of these frequency estimates to obtain the proportion of recombinant chromosomes in the gamete produced, *i. e.* the recombination fraction, and the frequencies of double reduction at the two loci employed. These three elements are directly obtainable from the frequency compounds open to estimation.

Although all four modes of gamete formation involving double reduction at one locus but not at the other are represented in the observed frequencies, the two modes of gamete formation involving double reduction at both loci have not yet certainly appeared, and the inference may be safely drawn that the centromere lies between the two loci used, as shown in Fig. 1.

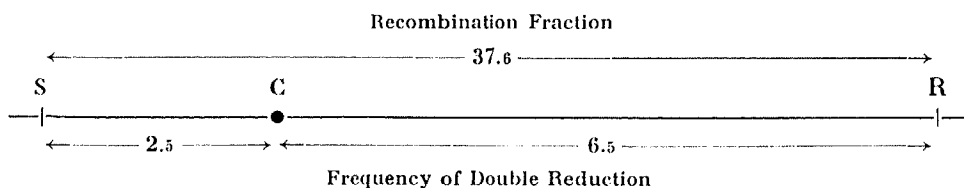


Fig. 1. The Location of the Genes for Short and Purple.

It should be observed that cases of bivalent formation will contribute to the recombination fraction, but that double reduction can only occur when a quadrivalent is formed at meiosis. The ratio of the recombination fraction to the frequency of double reduction therefore provides a rough indication of the frequency with which this homologous group of chromosomes forms bivalents or a quadrivalent.

LITERATURE CITED.

1. FISHER, R. A. 1947. The theory of linkage in polysomic inheritance. — *Phil. Trans. R. Soc. B.* No. 594. Vol. 233, pp. 55–87.

COMMENT

See Paper 262.