

SUPER-RECOMBINATION IN THE SEX CHROMOSOME OF THE MOUSE

A. R. G. OWEN
Department of Genetics, Cambridge

Received 1. ix. 52

I. INTRODUCTION

THE occurrence of recombination fractions significantly in excess of 50 per cent. may conveniently be named *super-recombination*. A case of this kind was reported (1) by Mrs M. E. Wallace (M. E. Wright) in 1947. She found that in the house mouse the two factors wv_2 ($wavy_2$) and sh_2 ($shaker_2$) showed respectively with *sex* the recombination values 56.07 and 56.73 per cent., but only 31.06 per cent. recombination with one another. At the same time a theoretical explanation (2) (principally due to Sir Ronald Fisher) was put forward which showed that these results could be accounted for on the basis of genetical interference. The theoretical treatment was based on the following assumptions.

- (a) An interference metric u exists on which the intercept lengths formed on a strand by consecutive exchange points are distributed independently in accordance with a frequency distribution

$$dp = f(u)du.$$

- (b) The interference function $f(u)$ approximates to

$$f(u) = \frac{d}{du}(-\operatorname{sech} \frac{1}{2}\mu u).$$

- (c) There is negligible interference across the centromere.

These assumptions are general in character and have been maintained in later work (3-8), except for a slight modification of the form of $f(u)$. At the time a further assumption was made relative to the situation in the sex chromosome of the mouse, namely that

- (d) the centromere is remote from the sex-determining portion.

In the later work on this theory (loc. cit.) it was found possible to take account of the fact that the arms of chromosomes are not infinitely long in comparison with the mean intercept length but have termini at finite distances. I have given (4) a brief account of this part of the theory, and a more extended account will be published elsewhere. In 1948 I carried out a further discussion of Wright's results in order to consider them with reference to finite arm length. The work was much facilitated by the fact that it had proved possible to replace the form for $f(u)$ quoted above by the simpler one $4ue^{-2u}$, which has been found to give predictions in agreement with Kosambi's theory (of proved utility (9-11)) and with data on interference in *Drosophila melanogaster* (8, 12). This discussion was not published at

the time, though it appeared as part of a Ph.D. dissertation (12). I think, however, that now it may be of some interest since Carter and Robertson in a recent paper (13) have drawn attention again to Wright's data as an apparently well substantiated case exhibiting super recombination due to loose linkage and genetical interference. Elsewhere (10) I have made criticisms of Carter and Robertson's particular modification of recombination theory, and in this note wish only to consider those of their theses which are relevant to the interpretation of Wright's data. They say that when all chiasma pairs involve four strands, and chiasma interference is absent, then the maximum recombination fraction is 53.34 per cent. This is equivalent to the statement that with the function $4ue^{-2u}$ recombination cannot exceed 53.34 per cent. This is indeed true for an infinitely long chromosome arm, which is the only model studied by Carter and Robertson, but is not true when account is taken of arm lengths that though long are less than infinite. It will be seen from the present discussion that recombinations of the order of 60 per cent. are possible with the function $4ue^{-2u}$ when allowance is made for the arm terminus.

There is a second point, also relevant, which merits some examination. Carter and Robertson correctly point out that super-recombination requires some degree of chromatid interference, and is not incompatible with there being some degree of chiasma interference. However, they imply that the function $4ue^{-2u}$ is of limited applicability, its use precluding the assumption of any chiasma interference, and in particular is not competent to explain Wright's data. It is indeed true that this function applies exactly only when chiasma interference is absent and chromatid interference is complete. However, I showed in 1948 (12) that it applies to a good degree of approximation (which is all that is necessary in view of the relative imprecision of even the best genetical data) in a variety of other circumstances, *i.e.* with various degrees of incomplete chromatid interference allied with appropriate intensities of chiasma interference. Its use for practical purposes seems therefore to be admissible on ample theoretical grounds, quite apart from the empirical support given by Kosambi's formula (*loc. cit.*) and data on *Drosophila melanogaster* (*loc. cit.*).

I find the following example particularly striking though it is only one of a complete sequence that can be constructed. Let us assume a degree of chiasma interference such that, relative to a genetical interference metric u , the *chiasma* intercept length is distributed as

$$dp = ue^{-u}du,$$

and suppose that chromatid interference operates in such a way that when a strand has been involved in a chiasma it has a probability of 1/3 (as opposed to 1/2) of being involved in the next chiasma. Then it will be found that the genetical interference function is

$$g(u) = 16e^{-4u} \left(\frac{1}{\sqrt{3}} \sinh \frac{4u}{\sqrt{3}} - u \right).$$

The mathematical form of $g(u)$ has indeed but a small resemblance to the more modest $f(u) = 4ue^{-2u}$ but tabulation shows that the two functions closely mimic one another over the range ($0 \leq u \leq 3$) *i.e.* over about 275 cM of map distance. Table 1 gives their value together with the absolute discrepancy (error column) and the discrepancy as percentage of the value of $f(u)$. The error is trivial until u exceeds 3. Beyond this point it might indeed affect delicate phenomena such as super-recombination on arms longer than 275 cM. For shorter arms, however, the divergence at the tails is irrelevant, and the functions are to all intents and purposes the same.

TABLE 1

Metric u	Map length x cM	$f(u)$	$g(u)$	Error	Per cent. error
0	0	0	0	0	...
0.25	9	0.60653	0.60242	-0.00411	-0.68
0.50	28	0.73596	0.70483	-0.03113	-4.23
0.75	50	0.66939	0.66251	-0.00688	-1.02
1.00	75	0.54134	0.55101	0.00967	1.79
1.50	125	0.29872	0.30635	0.00763	2.55
2.00	175	0.14653	0.14655	0.00002	0.00
2.50	225	0.06738	0.06575	-0.00163	-2.42
3.00	275	0.02975	0.02873	-0.00102	-3.43
4.00	375	0.00336	0.00528	0.00192	57.14

In showing that the function $4ue^{-2u}$ could be regarded as having its genesis in an approximation to special models involving explicitly postulated mechanisms of chromatid and chiasma interference, I do not wish to imply that there is any warrant for putting forward such detailed models as effective descriptions of reality. I believe that the actual phenomena are in fact so complicated as to be outside the scope of manageable mathematical description. It seems much the more realist policy to regard genetical interference as a unitary phenomenon (which probably it is) described adequately by a specification of the exchange point distribution on any one of the meiotic strands. This self denying ordinance keeps us within the range of concepts capable of some verification from genetical observations, *i.e.* the breeding behaviour of organisms.

To sum up, we can say that use of $4ue^{-2u}$ involves us in no patent subservience to over-restricted models of crossing-over, nor in any patent conflict with theories of crossing-over. Its use is consistent with various postulated partitions of interference between chromatid and chiasms components. When finite arm length is taken into account it is capable of yielding values exceeding 53 per cent., should such values be required by experimental data of sufficient precision. These remarks apply with equal force to any close mimic of the function *e.g.* to

$$\frac{d}{du}(-\operatorname{sech} \frac{1}{2}\pi u).$$

2. ASSUMPTIONS REGARDING THE SEX CHROMOSOME

In the male no chiasmata are likely to occur in the differential segment. In the original discussion we therefore took P, the point of junction of the pairing segment and the differential segment, as the effective terminus of one arm of the sex-chromosome. Since there were no chiasmata in the differential segment beyond P we assumed that the exchange point E formed nearest to P was exempt from any interference originating between it and P or beyond P. Thus the probability distribution of u , the metrical distance of this exchange point from P, was chosen to be a function of Poisson type, namely

$$\frac{d}{du}(-e^{-\frac{1}{2}\pi u}),$$

which is the limiting form of the tail of the distribution

$$\frac{d}{du}(-\operatorname{sech} \frac{1}{2}\pi u).$$

The centromere was assumed to be effectively at an infinite distance from P, so that the only repressive influences inhibiting the formation of exchange points further from P than E were exerted by established exchange points and represented by the interference function

$$\frac{d}{du}(-\operatorname{sech} \frac{1}{2}\pi u).$$

The map distance and recombination between P and a locus at metrical distance t from P were then tabulated numerically.

The assumptions about the terminus P were in fact equivalent to those used by the writer (4) in an analysis appropriate to arms of finite length. The numerical study may therefore be repeated using the function $4ue^{-2u}$ and the established formula (loc. cit.).

3. CENTROMERE ASSUMED REMOTE

Retaining the hypothesis that the arm is of very great length and the centromere very remote from P, let $x(t)$ and $y(t)$ be the map distance and recombination between the terminus and a locus at metrical distance t from it. Then, with the function $4ue^{-2u}$, we have

$$\begin{aligned} x(t) &= t + \frac{1}{4} - \frac{1}{4}e^{-4t} \\ y(t) &= \frac{1}{2} - \frac{1}{2}e^{-2t}(\cos 2t - \sin 2t). \end{aligned}$$

These functions are tabulated in table 2, along with the corresponding figures for the function

$$\frac{d}{du}\left(-\operatorname{sech} \frac{\pi u}{2}\right)$$

transcribed from the earlier paper (2).

The new function gives a higher maximum for y than does the earlier one. The new value is 60.394 per cent., *i.e.* $\frac{1}{2}(1 + e^{-\frac{1}{2}\pi})$, as opposed to 55.8 per cent. The points where y has values 56 and 57 per cent. are more than 60 cM apart, a fact which somewhat strains agreement with the known separation of wv_2 and sh_2 , of the order 35-40 cM only.

TABLE 2

Recombination and map distance between end of pairing segment and locus at metrical distance t when the centromere is infinitely remote

t	Old function		New function	
	$x(t)$ cM	$y(t)$ per cent.	$x(t)$ cM	$y(t)$ per cent.
0	0	0	0	0
0.1	14.598	14.474	18.242	18.013
0.2	27.430	26.496	33.767	32.182
0.3	39.069	36.125	47.470	43.082
0.4	49.955	43.497	59.953	50.464
0.5	60.399	48.831	71.616	55.540
0.6	70.601	52.405	82.707	58.579
0.7	80.687	54.535	93.480	60.055
0.8	90.727	55.734	103.981	60.361
0.9	100.756	55.734	114.317	59.927
1.0	110.789	55.275	124.542	58.969
1.1	120.832	54.688	134.693	57.740
1.2	130.886	53.846	144.794	56.409
1.3	154.862	55.097
1.4	164.908	53.884

Thus the quantitative difference between the old and the new functions, while very slight, shows up in the rather delicate phenomenon of super-recombination. The functions are such close mimics of one another that their predictions in respect of the short-range properties of linkage groups are not readily distinguishable. They are similarly equally compatible with *a priori* theoretical requirements as considered above. The discrepancy in the present application does however show how unwise it would be to dogmatise about numerical magnitudes in the more delicate long-range phenomena of which super-recombination is one.

4. ADMISSABLE HYPOTHESES WHEN THE CENTROMERE IS ASSUMED LESS REMOTE

The values of the recombinations of wv_2 and sh_2 with sex require both loci to be fairly remote from the terminus (90 cM or more), unless the interference is much in excess of the Kosambi level *i.e.* that associated with $4ue^{-2u}$, (*cf.* Owen, 1951, (6)). With interference of the Kosambi order of magnitude, wv_2 and sh_2 must both be 90 cM

or more from the terminus but the observed values are compatible with the centromere being located either between or beyond them.

Table 3 shows the maximum recombination value with the terminus which a locus on arms of various lengths may attain, when the interference function is $4ue^{-2u}$. The recombination between

TABLE 3

Maximum recombination with terminus (y_m) and recombination between terminus and centromere (y_0) on arm of metrical length T

T	Arm length cM	y_m per cent.	y_0 per cent.
0.785	72.033	50.000	50.000
0.800	73.734	50.582	50.566
0.900	85.213	53.569	53.656
1.000	96.403	55.892	55.531
1.100	107.331	57.361	56.442
1.200	118.040	58.356	56.635
1.300	128.574	59.027	56.330
1.400	138.968	59.479	55.710
1.500	149.258	59.782	54.917
1.600	159.469	59.986	54.062
1.700	169.621	60.123	53.223
1.800	179.732	60.215	52.245
1.900	189.810	60.277	51.769
2.000	199.866	60.318	51.197
∞	∞	60.394	50.000

centromere and terminus is also indicated. For arms of the order of 200 cM the maximum value y_m is close to the value achieved on the infinite arm. This will be equally true when the function is

$$\frac{d}{du}(-\operatorname{sech} \pi u/2).$$

Hence the original treatment will apply to a good order of approximation if the arm is of this length.

To ascertain what arm length (T in metrical units) is compatible with the hypothesis $f(u) = 4ue^{-2u}$, the course of the recombination fraction has been plotted for $T = 1.40$ and $T = 1.00$. In the former case the centromere would lie beyond wv_2 and sh_2 at 138.97 cM from the end of the differential segment. x and y are shown in table 4. Close agreement with Wright's data is not obtained. The maximum of y is too large on this theory, while wv_2 and sh_2 would be required to be about 50 cM apart.

Table 5 shows x and y calculated on the hypothesis that the centromere lies at $T = 1.00$, and at 96.4 cM from the differential segment. It is assumed that there is no interference across the

centromere. The agreement is about as good as that obtained in the original treatment.

The last hypothesis to be examined is that the centromere lies between the pairing segment and the factors wv_2 , sh_2 . This is easily

TABLE 4
*Map length and recombination over a terminal segment
or an arm of metrical length 1.40*

t	$x(t)$ cM	$y(t)$ per cent.
0	0	0
0.1	18.09	17.865
0.2	33.45	31.887
0.3	46.97	42.336
0.4	59.19	49.882
0.5	70.58	54.827
0.6	81.28	57.756
0.7	91.46	59.155
0.8	100.89	59.461
0.9	110.29	59.058
1.0	118.78	58.265
1.1	126.38	57.344
1.2	132.72	56.511
1.3	137.21	55.925
1.4	138.97	55.710

TABLE 5
*Map distance from and recombination with terminus when centromere
is at 96.64 cM from terminus*

t	$x(t)$ cM	$y(t)$ per cent.
0	0	0
0.1	17.513	17.293
0.2	32.249	30.744
0.3	45.033	40.097
0.4	56.373	47.625
0.5	66.556	52.062
0.6	75.654	54.560
0.7	83.595	55.660
0.8	90.091	55.876
0.9	94.635	55.686
1.0	96.403	55.531
1.1	98.170	53.618
1.2	102.715	52.130
1.3	109.211	51.029
1.4	117.152	50.263

verified to be quite incompatible with interference at the Kosambi level. Recombinations of the right order could only be obtained with interference of much greater ferocity, with a mean intensity of the order 0.8 (*cf.* Owen, 1949 (3)).

To summarise, it may be said that Wright's data are compatible with interference of Kosambi order if either

- (a) the arm length is of the order of 200 cM, with the centromere lying beyond wv_2 and sh_2 (the old interference function giving a better fit than the new), or
- (b) the centromere is situated between these factors and distant about 90-100 cM from the differential segment (the new interference function being used).

The data are definitely incompatible with the assumption that the centromere is near to the differential segment unless it is the case either that interference is operating with exceptional severity or that there is strong interference across the centromere.

5. SUMMARY

It is shown that contrary to the supposition of Carter and Robertson, the functions used to represent genetical interference are compatible with four strand crossing-over without complete chromatid interference and with some chiasma interference. It is also made clear that when account of the chromosome termini is taken these functions yield recombination values in excess of 53.34 per cent. and are competent to explain Wright's data on wv_2 , sh_2 and sex in the house mouse. It is shown that these data are consistent with a centromere at a distance of the order of 200 cM from the differential segment, with wv_2 and sh_2 lying between the centromere and the pairing segment, or with the centromere lying between these factors and at about 100 cM from the differential segment.

6. REFERENCES

- (1) WRIGHT, M. E. 1947. Two sex linkages in the house mouse, with unusual recombination values. *Heredity*, 1, 349-354.
- (2) FISHER, R. A., LYON, M. F., AND OWEN, A. R. G. 1947. The sex chromosome in the house mouse. *Heredity*, 1, 355-365.
- (3) OWEN, A. R. G. 1949. The theory of genetical recombination. I. Long-chromosome arms. *P.R.S., B*, 136, 67-94.
- (4) OWEN, A. R. G. 1950. The theory of genetical recombination. *Advances in Genetics*, 3, 117-157.
- (5) FISHER, R. A. 1951. A combinatorial formulation of multiple linkage tests. *Nature*, 167, 520.
- (6) OWEN, A. R. G. 1951. An extension of Kosambi's formula. *Nature*, 168, 208.
- (7) OWEN, A. R. G. 1953. The analysis of multiple linkage data. *Heredity*, 7, 000.
- (8) OWEN, A. R. G. Genetical interference in *Drosophila melanogaster*. (not yet published).
- (9) BHAT, N. R. 1948. An improved genetical map of Punnett's "B" chromosome in the Sweet Pea, *Lathyrus odoratus* L. *J. Genet.*, 48, 343-358.
- (10) BHAT, N. R. 1950. A case in rice, *Oryza sativa* L., supporting Kosambi's formula for estimating map lengths. *Amer. Nat.*, 84, 71-80.
- (11) BORGER, R. 1950. The order of genes in the fifth linkage group of the house mouse. *Nature*, 166, 697.
- (12) OWEN, A. R. G. 1948. *Ph.D. Dissertation*. University Library, Cambridge, pp. 172-228.
- (13) CARTER, T. G., AND ROBERTSON, A. 1952. A mathematical treatment of genetical recombination using a four-strand model. *P.R.S., B*, 139, 410-426.