THE SEX CHROMOSOME IN THE HOUSE MOUSE

R. A. FISHER, M. F. LYON and A. R. G. OWEN
Department of Genetics, University of Cambridge

Received 5.vi.47

I. THE FORMULATION OF LINKAGE RELATIONSHIPS

SINCE Morgan's [7] original proposal in 1911 of a linear arrangement for genes in the same chromosome, it has been obvious that it should be possible to specify, in terms of the lengths, in some appropriate metric, of the intercepts between them, the entire series of gametic frequencies for an organism heterozygous at any number of linked loci. The lack of such a formal specification constitutes the most serious gap in current expositions of genetical theory. The following attempt at such a formulation, stimulated by the experimental data recently obtained in this department on the sex chromosome of the house mouse (Wright,[4] 1947), owes much to earlier writers, among whom especially should be mentioned K. Mather [1] (1937) and Kosambi [2] (1944), both of whom have supplied ideas essential to what follows. The comparisons of the theoretical expectations with the relevant observations in other cases is a task beyond the scope of the present paper.

Any segment of a chromosome bounded by two marked loci, which enters a gamete, will have experienced in the preceding meiosis 0, 1, 2, . . . breaks involving interchange with one or other strand from the homologous chromosome. These occurrences are mutually exclusive and exhaustive, so that if p_0, p_1, p_2, . . . represent the frequencies of these events, it follows that

$$p_0+p_1+p_2+ \ . \ . \ . \ = 1.$$

With these quantities we may associate others, s_r, representing the probabilities of r or more interchanges, so that

$$s_r = \sum_{t=r}^{\infty} p_t.$$

We may recognise in s_r the probability integral of the position of the r^{th} break from the first locus, the integral being evaluated at the position of the second locus. The values s_r like p_r are always positive, and each value is less than the last by the deduction of the corresponding p, i.e.,

$$s_r - s_{r+1} = p_r.$$

 Heredity, 1: 355-365, (1947).

The observable recombination fraction between the two loci, denoted by y, is seen to be

$$\begin{aligned} y &= p_1 + \quad p_3 + \quad p_5 + \ \ldots, \\ &= s_1 - s_2 + s_3 - s_4 + s_5 - \ \ldots, \end{aligned}$$

while the map distance, representing the average number of interchanges in the segment chosen is

$$\begin{aligned} x &= p_1 + 2p_2 + 3p_3 + 4p_4 + \ \ldots, \\ &= s_1 + \ s_2 + \ s_3 + \ s_4 + \ \ldots. \end{aligned}$$

The latter quantity is additive, since if B lies between A and C, the average number of interchanges between A and C must be the sum of the average numbers between A and B and between B and C. Consequently x, though not directly observable, supplies a consistent metric for mapping.

Kosambi has shown that, to a very satisfactory approximation in many cases, the two genetically important quantities x and y are connected by the relation,

$$\tanh\,(2x) = 2y,$$

This relation suffices to specify the gametic series of heterozygotes at three loci, but not at four or more, for which we should need expressions for the entire series $s_1, s_2, s_3, \ldots$ in terms of some common parameter, in terms of which, therefore, both x and y could be specified.

Owing to the probability of disturbances in the neighbourhood of the centromere, and of the ends of the chromosome, we shall not assume the exactitude of Kosambi's relationship, but instead shall introduce a metric u, in terms of which interference of neighbouring interchanges shall be uniform, and in particular for which the probability of an interchange is reduced in the ratio

$$\tanh\,(\tfrac{1}{2}\pi u) \qquad (1)$$

by the influence of an interchange already established at a distance u.

Since

$$\int \tanh\,\phi\, d\phi$$

is

$$\log\cosh\,\phi,$$

and

$$e^{-\log\cosh\,\phi}$$

is

$$\operatorname{sech}\,\phi,$$

it follows that the probability that the next interchange to the one already established lies in the intercept du is

$$\begin{aligned} &\operatorname{sech}\,(\tfrac{1}{2}\pi u)\,\tanh\,(\tfrac{1}{2}\pi u)\,d(\tfrac{1}{2}\pi u) \\ &\quad = d(-\operatorname{sech}\,\tfrac{1}{2}\pi u). \end{aligned} \qquad (2)$$

This expression therefore defines the frequency distribution of the length of intercept between two adjacent interchanges. The theory we shall develop postulates that this distribution is independent of all

more remote breaks, so that the lengths of adjacent intercepts are distributed independently.

The average value of the intercept length, u, found by evaluating

$$\int_0^\infty u d(-\operatorname{sech} \tfrac{1}{2}\pi u)$$

is unity, since by partial integration this expression is equal to

$$\int_0^\infty \operatorname{sech} \tfrac{1}{2}\pi u \, du,$$

or to

$$-\frac{4}{\pi}\int_0^\infty \frac{d(e^{-\frac{1}{2}\pi u})}{1+(e^{-\frac{1}{2}\pi u})^2};$$

this becomes

$$\frac{4}{\pi}\int_0^1 \frac{dz}{1+z^2} = \frac{4}{\pi}\Big[\tan^{-1} z\Big]_0^1 = 1,$$

on making the substitution

$$z = e^{-\frac{1}{2}\pi u}.$$

It is for this reason that the argument $\frac{1}{2}\pi u$ has been used, rather than any other multiple of u; for in regions far from either end of a chromo-

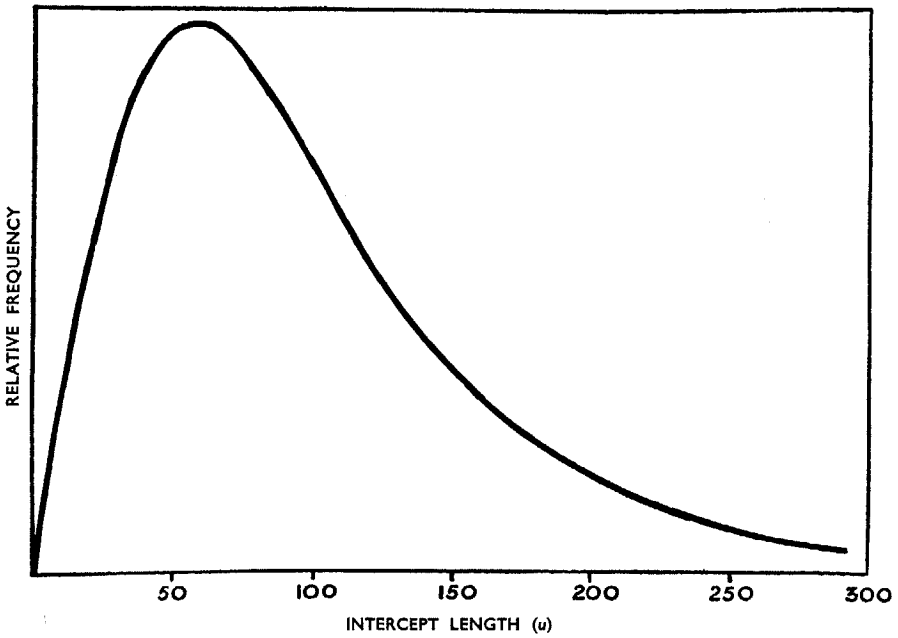

FIG. 1.—Frequency distribution of intercept length sech $(\frac{1}{2}\pi u)$ tanh $(\frac{1}{2}\pi u)$ $d(\frac{1}{2}\pi u)$ adopted as basis for calculation.

some arm, this convention leads us to expect the average density of breaks to be unity, so that u will closely simulate the map distance in such regions. The distribution of intercept length is shown in fig. 1.

2. THE PAIRING SEGMENT OF THE SEX CHROMOSOME

Let us now consider the special case afforded by the pairing segment of the sex chromosome in an organism such that (as Koller [3] and Darlington (1934) have shown in *Rattus norwegicus*) chiasmata may occur between the centromere and the sex-determining portion. In the male no chiasmata can occur in the differential segment. Consequently, the repressive influence of neighbouring breaks will be reduced at the adjacent end of the pairing segment, and qualitatively we might expect the density of breaks to be raised in this region.

Arguing more exactly, we may say that the distribution of the first break in the pairing segment will be that of a value u which is constrained to exceed some arbitrarily large value U. In other words, the distribution will be that of the limiting form of the tail of the original distribution, or

$$df = e^{-\frac{1}{2}\pi u_1}\, d(\tfrac{1}{2}\pi u_1),$$

where u_1 is now measured from the end of the pairing segment.

The probability that there will be at least one break within a distance u from the end of the pairing section is now given by

$$1-p_0 = s_1 = \int_0^u e^{-\frac{1}{2}\pi u_1}\, d(\tfrac{1}{2}\pi u_1)$$

so that

$$p_0 = e^{-\frac{1}{2}\pi u}$$
$$s_1 = 1-e^{-\frac{1}{2}\pi u}.$$

The probability of exactly one break may be evaluated by integrating, for values of u_1 less than u, the expression

$$e^{-\frac{1}{2}\pi u_1}\, d(\tfrac{1}{2}\pi u_1)\ \operatorname{sech}\ \tfrac{1}{2}\pi(u-u_1),$$

which is the product of the probability of a first break in the interval du_1 by the probability that there shall be no further break within the length u from the terminus.

This probability is, therefore,

$$p_1 = \int_0^u e^{-\frac{1}{2}\pi u_1}\ \operatorname{sech}\ \tfrac{1}{2}\pi(u-u_1)\, d(\tfrac{1}{2}\pi u_1).$$

By making the successive substitutions

$$v = \tfrac{1}{2}\pi(u-u_1),\ z = e^v,$$

its value is found to be

$$e^{-\frac{1}{2}\pi u}\int_0^{\frac{1}{2}\pi u} e^v \operatorname{sech}\ v\ dv = e^{-\frac{1}{2}\pi u}\int_1^{e^{\frac{1}{2}\pi u}} \frac{2z\ dz}{1+z^2}$$

$$= e^{-\frac{1}{2}\pi u}\Big[\log\ (1+z^2)\Big]_1^{e^{\frac{1}{2}\pi u}} = e^{-\frac{1}{2}\pi u}\log\frac{1+e^{\pi u}}{2}$$

so that

$$p_1 = e^{-\frac{1}{2}\pi u}\{\log\cosh\tfrac{1}{2}\pi u+\tfrac{1}{2}\pi u\}$$

and

$$s_2 = s_1-p_1 = 1-e^{-\frac{1}{2}\pi u}\{\log\cosh\tfrac{1}{2}\pi u+\tfrac{1}{2}\pi u+1\}.$$

For breaks subsequent to the second the analytic forms are difficult, and recourse may be had to mechanical integration. Characteristic values of s for the distribution of the first few breaks are given in table 1. Fig. 2 shows the corresponding frequency distributions.

TABLE 1

Probabilities of at least 1, 2, 3, . . . breaks between given points and the sex-determining segment, with corresponding values of the map distance (x), *and the recombination fraction* (y).

Distance in metric chosen u	s_1	s_2	s_3	s_4	s_5	s_6	x per cent.	y per cent.
0·00	0	0	0	0	0	0	0	0
0·10	·14536	62					14·598	14·474
0·20	·26960	467	3				27·430	26·496
0·30	·37577	1472	20				39·069	36·125
0·40	·46651	3228	76	1			49·955	43·497
0·50	·54406	5780	209	4			60·399	48·831
0·60	·61034	9086	469	12			70·601	52·405
0·70	·66698	·13044	912	32	1		80·687	54·535
0·80	·71539	·17520	1594	72	2		90·727	55·734
0·90	·75676	·22364	2564	147	5		100·756	55·734
1·00	·79212	·27434	3859	273	11		110·789	55·275
1·10	·82234	·32601	5502	470	24	1	120·832	54·688
1·20	·84816	·37754	7503	763	46	2	130·886	53·846

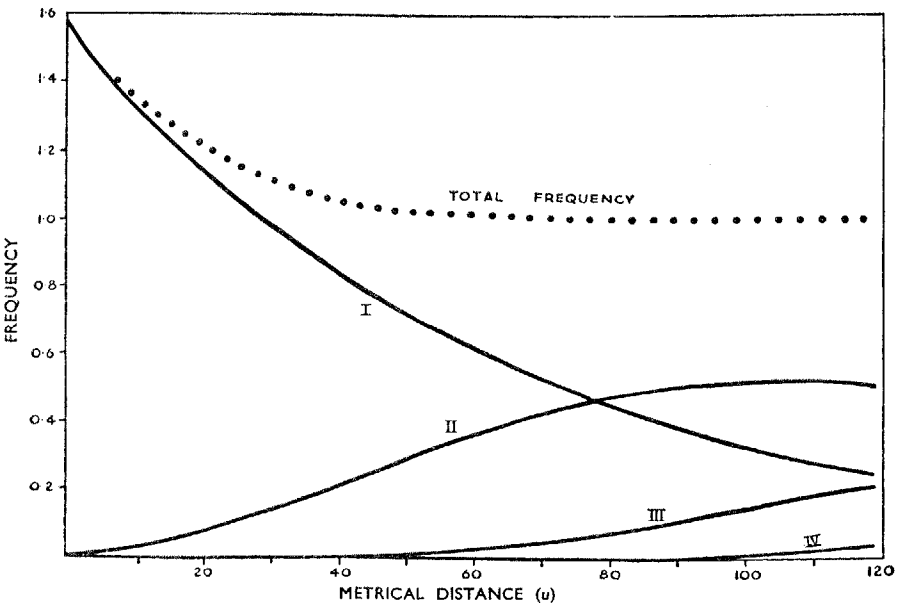

FIG. 2.—Distributions of the first four breaks, counting from the junction of the differential and pairing segments, in terms of the metrical distance (u) from this junction.

The sum of the frequency ordinates for these successive breaks gives the total density of interchange in relation to the metric u adopted. It will be seen that this density is slightly enhanced in the terminal region, but that it soon becomes almost uniform, so that increases of u become very nearly equivalent to equal increases in map distance.

We are now in a position to compare the map distance measured from the sex-determining portion of the chromosome, with the recombination fraction observed between a given marker gene and sex. The comparison is shown in fig. 3. It will be seen that the

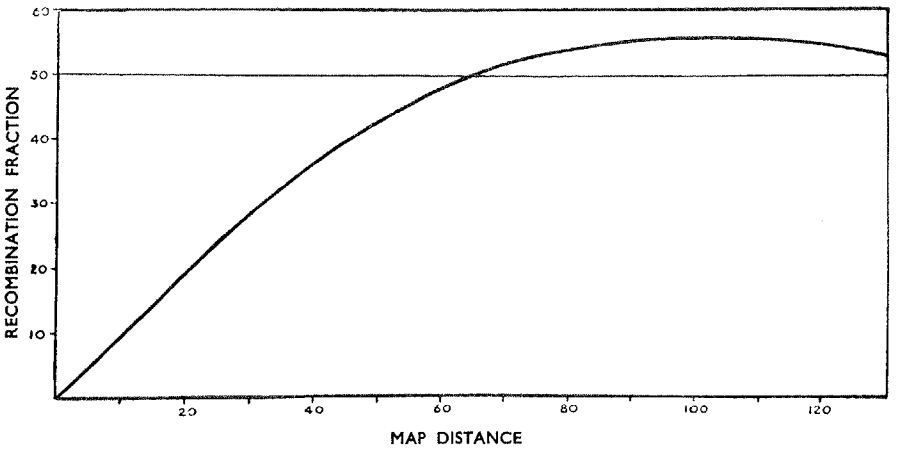

FIG. 3.—Relation between observable recombination (y) with sex and map distance (x) from the differential segment.

recombination fraction, rising initially in equality with the map distance, attains a maximum of about 55·75 per cent. at a map distance of 98 or 99 centimorgans, and thereafter falls gradually towards 50 per cent. For a considerable distance of more than 30 centimorgans the recombination fraction exceeds 55 per cent., and can be detected by reasonably careful experimentation. At positions from about 58 to 68 centimorgans from the terminus the difference from 50 per cent. recombination will be almost imperceptible, and genes in this region will appear to be inherited almost independently from sex, as will be the case also at distances much greater than we have explored.

For lengths much greater than our figures illustrate, it is to be presumed that the calculations will become unrealistic, owing to the unknown position of the centromere, which in these figures has been taken to be infinitely distant. They should, therefore, be regarded as provisional and subject to revision in the light of a more exact examination. Qualitatively, however, we should expect them to give a good representation of the genetically observable relations of sex-linked genes.

3. COMPARISON WITH THE LINKAGE RELATIONSHIPS OF THREE GENES IN THE HOUSE MOUSE

During 1945 Wright, working in this department, observed in several lines carrying wv_2 or sh_2 a linkage-like disturbance of the sex ratio, and proceeded to set up the definitive test reported by her in this number.[4] In the following year Falconer observed in Line 9 of these stocks a rather close linkage between the dominant waving gene Rex (*Re*) and sh_2. Since *Re* and wv_2 have similar effects on the hair, they could not be tested in the same line, but a few special matings, which Falconer [5] has reported, were sufficient to show that Rex and wavy were not allelomorphic, but on the contrary that they are located on opposite sides of the shaker locus.

Snell and Law (1939) who first reported [6] the linkage of wv_2 and sh_2, give the recombination fraction as 25 per cent. in females, there being 64 recombinants among 256 young. For males they report only 56 young, and do not state whether their doubly heterozygous parents received the mutant or the normal genes from their fathers. These data cannot therefore be used for examining sex linkage, but give an estimate of 26·8 per cent. for recombination between wv_2 and sh_2 in male gametogenesis. We are indebted to Dr Snell, also, for some early intercross data (table 2) in which the males are known to have received both mutant genes from the dam.

TABLE 2

	++	wv_2	sh_2	wv_2sh_2	Total	Per cent. of expectation	
						wv_2	sh_2
♀	110	13	19	12	154	64·9	80·5
♂	117	22	16	16	171	88·9	74·9
Total . . .	227	35	35	28	325	...	...

So far as these go they give maximum likelihood estimates of 56·37 per cent. for the recombination of wavy and sex, in confirmation of Wright's value of 56·07 per cent., but of 47·65 per cent. for sex and shaker, which is significantly lower than Wright's value, 56·73 per cent. Apart, however, from being intercross data in which accurate allowance for differential viability is not practicable, it should be noticed that both recessives are seriously below expectation in frequency. The corresponding estimates by the product method are 57·24 per cent. for wavy and sex, and 48·30 per cent. for shaker and sex. Even the product method, however, can scarcely in this case eliminate the viability disturbance, since the deficiency may not improbably be unequal in the two sexes. Indeed, the average recombination found

here, 52·77 per cent., might be interpreted to exceed 50 per cent. only by reason of a greater elimination of recessive females. On the whole it seems best to use Wright's values, without attempting to adjust them by the use of these earlier observations.

Somewhat clearer confirmation is shown by some more recent backcross data, which Dr Snell has since kindly added. These are shown in table 2*a*.

TABLE 2*a*

$$\frac{wv_2sh_2X}{wv_2sh_2X}\,♀ \times \frac{++Y}{wv_2sh_2X}\,♂$$

	++	wv_2	sh_2	wv_2sh_2	Total
♀	27	8	13	12	60
♂	14	5	7	15	41
Total	41	13	20	27	101

$$\frac{wv_2sh_2X}{wv_2sh_2X}\,♀ \times \frac{wv_2sh_2Y}{++X}\,♂$$

	++	wv_2	sh_2	wv_2sh_2	Total
♀	13	1	5	6	25
♂	9	1	3	4	17
Total	22	2	8	10	42

There is again some abnormality in the single factor ratios, but so far as recombination is concerned these mice, 143 in all, give estimates very close to those of Wright's experiment. For linkage of wavy with sex we have here 79/143, or 55·245 per cent., for shaker with sex 78/143, or 54·545 per cent., and for wavy with shaker 43/143, or 30·070 per cent., all in close agreement with the values found in our own department.

The paradoxical nature of Wright's data consists in two points : (i) the appearance of recombination fractions between sex and wavy and between sex and shaker both significantly exceeding 50 per cent. ; (ii) the fact that two loci distant from each other by about 35 units of map distance should be almost equally related, as regards recombination, with sex. With three-point backcross data we are accustomed to recognise the order of the genes by the doubly recombinant class being distinctly rarer than all others. In these data, however, the striking and most significant difference lies between the old combinations, and those in which both wavy and shaker are separated from sex.

These remarkable features are perfectly in accordance with

expectations based on fig. 3. We must suppose that both the loci for wv_2 and for sh_2 lie near the maximum, probably on either side of it. There is apparently plenty of room in this region for loci 35 or 40 units apart having recombination fractions insignificantly different from those observed. The data do not by themselves suffice to determine which of the two is nearer to the sex-determining region of the chromosome.

The order is, however, rendered very probable by the evidence concerning Rex. The locus of Rex is estimated, on the small body of data available to Falconer, to be about 20 units from shaker, on the side distant from wavy. A summary of progenies from Rex males observed in this department gives to date the following frequencies (table 3).

TABLE 3

	+♀	+♂	*Re*♀	*Re*♂	Total
Coupling . . .	119	108	116	120	463
Repulsion . . .	64	59	46	57	226

The ratios of recombinants to old combinations are therefore 224 : 239 in coupling, and 121 : 105 in repulsion. It is obvious that recombination does not differ significantly from 50 per cent., and this indicates that the locus of Rex must be about 64 units from the end of the map. A more careful estimate, eliminating the possibility of small differences between the relative viabilities of Rex animals in the two cases, is supplied by the equation

$$\frac{y}{1-y} = \sqrt{\frac{224}{239} \cdot \frac{121}{105}}$$

from which it appears that y, estimated from these data, is about 50·96 per cent.

It is therefore easy to assign positions to these three loci consistent with the whole of the genetical information so far available about them. With more data a closer tie-up should be possible, and it may even be necessary to consider such observable effects as the centromere of the sex chromosome may possibly have on these relationships. Triple backcrosses of males with wv_2 and sh_2 females from independent stock, and with sh_2 and *Re* are already in hand. As a provisional map, however, we propose

Locus	Map distance from terminus
Re	65 cm.
sh_2	85 ,,
wv_2	120 ,,

It is still possible, however, that these three loci may lie in the reverse order with Rex some way beyond the centromere.

DISCUSSION

The genetic phenomena of the sex chromosome in mice display the two surprising features, (i) of recombination fractions significantly exceeding 50 per cent., and (ii) of two loci, those of wavy and shaker, separated by a considerable interval of about 35 map units, showing no apparent difference in their linkage with a third locus, namely, that of sex, to be identified with the end of the pairing segment. Even though these phenomena are in harmony with the theory of interference which we have put forward, and which contains no strikingly novel feature, it may well be asked why such phenomena have not been observed in other organisms.

If the basis of our calculations is at least approximately valid, the conditions for the observation of recombination fractions exceeding 50 per cent. appear to be (*a*) the use of a marker gene located near the end of the chromosome, and (*b*) that the chromosome arm on which it is situated should be of considerable map length. The first of these conditions is admirably met by using sex itself as a marker, and probably is rarely satisfied by the location of autosomal genes.

The other organisms in which the pairing segment of the sex chromosome can be studied do not seem to provide the conditions for finding a similar genetic situation. In *Lebistes reticulatus*, where partial sex linkage, or inheritance in the pairing segment, was first demonstrated, the known genes, though numerous, are all very closely linked with the sex-determining portion. They are, moreover, functionally associated with sex, being epigamic in character and sex-limited in manifestation to the male. Only if numerous genes of other kinds were available in this species could it be determined whether any were loosely linked with sex and, if so, whether in any such case the recombination fraction exceeded 50 per cent.

In Man the conditions for the detection of incomplete sex linkage are such that only relatively close linkage, up to about 30 per cent., can be recognised, so that, if any part of the sex chromosome were to show recombination with sex exceeding 50 per cent., but less than 70 per cent., it would, with the kind of data hitherto available, or likely to be available in the future, certainly appear to be inherited independently of sex.

In Drosophila both X-borne and autosomal genes are known located close to the ends of their respective chromosome arms, and these arms are not all so short genetically as to preclude absolutely recombination between the gene and the centromere slightly exceeding 50 per cent., though there is not much room for this. The bulk of genetic tests carried out with this genus, with a view to accurate mapping, has, however, for obvious reasons, been principally concerned with determining the shorter map intervals with precision with a view to inferring the longer map intervals from these by addition. Whether a search of the published data for large-scale

tests with longer intervals would yield any situation comparable to that which appears to prevail in the sex chromosome of the House Mouse we do not know, although it will clearly be profitable to try out our formulation of linkage relationships on the best of this material, and on other species where, with fewer genes observable, longer intervals have often been determined with care.

If our views are correct, the situation in the House Mouse, though qualitatively similar to that revealed by Koller and Darlington in the Rat, must differ quantitatively from that species in that the arm of the pairing segment bearing the sex-determining tract must, so far as we can judge, extend to some 80 units of map distance in the House Mouse ; whereas the frequency of equational separation of the centromeres observed by Koller and Darlington suggest that the distance in the Rat is not more than 5 to 10 map units. Further cytological studies of the Mouse will therefore be awaited with great interest.

REFERENCES

[1] MATHER, K. 1937.
The determination of position in crossing-over. II. The chromosome length-chiasma frequency relation.
Cytologia Fujii Jub. Vol. 514-526.

[2] KOSAMBI, D. D. 1944.
The estimation of map distances from recombination values.
Ann. Eug. 12, 172-175.

[3] KOLLER, P. C., and DARLINGTON, C. D. 1934.
The genetical and mechanical properties of the sex chromosome. I *Rattus norvegicus* ♂
J. Genet. 29, 159-173.

[4] WRIGHT, M. E. 1947.
Two sex linkages in the house mouse, with unusual recombination values.
Heredity 1, 349-354.

[5] FALCONER, D. S. 1947.
Linkage of Rex with shaker-2 in the house mouse.
Heredity 1, 133-135.

[6] SNELL, G. D., and LAW, L. W. 1939.
A linkage between shaker-2 and wavy-2 in the house mouse.
J. Hered. 30, 447.

[7] MORGAN, T. H. 1911.
Chromosomes and associative inheritance.
Science 34, 636-638.