Genet. Res., Camb. (1976) **28**, *pp.* 291–299
With 2 text figures
Printed in Great Britain

Distribution of crossing-over in mouse chromosomes

By MARY F. LYON
MRC Radiobiology Unit, Harwell, Oxon, OX11 0RD, England

(*Received* 2 *August* 1976)

SUMMARY

In devising maps of the positions of gene loci along chromosomes, by measurements of crossing-over, it is usual to assume that the chiasmata which give rise to the crossing-over are randomly located, unless there is evidence to the contrary. If chiasmata are in fact preferentially located at certain points or regions this not only affects mapping, but also is of interest for population genetics since it may favour the development during evolution of blocks of genes relatively undisturbed by crossing-over. Preferential localization of chiasmata is well known in a number of disparate organisms (Darlington 1965; John & Lewis, 1965), but has so far been little studied in mammals. Evidence is presented suggesting that chiasmata are in fact non-randomly located in certain chromosomes of the mouse.

1. METHODS AND RESULTS

Green (1975) has produced a map of the distribution of genes in linkage groups or chromosomes of the mouse, derived from published data on recombination (Fig. 1). Clearly the known and located genes are only a small fraction of those which must actually exist. However, one would suppose that the known genes would constitute a fairly random sample of all genes, at least in their location. They might well of course be non-random in other respects, such as mutation rate, since one would expect genes with high mutation rate to be discovered first.

Therefore, if (and only if) the known genes are randomly located, and chiasmata are also randomly located, one would expect the apparent positions of the genes on the linkage map also to be random. However, if the chiasmata were not random, but occurred preferentially in certain regions, one would expect to find apparent wide spacing of genes in the regions in which chiasmata occurred, and tight clusters of genes in other regions.

In order to test this point, Green's map was divided into numerous equal segments, by a series of parallel lines, at intervals of 1 cm. Green used a distance of 2 cm to represent a map distance of 10 cM (centiMorgans) and hence our subdivisions each represented 5 cM. In order to place the whole set of lines, the first was drawn just below the knobs used by Green to depict the centromeres. The numbers of gene loci in each 5 cM segment of each chromosome were then counted. If the distance of the centromere from the located genes in a given

chromosome was not known (as in chromosomes 2, 4, 5, 7, 8 and 10) the empty segments between the centromere and the first located gene were discounted. The last segment counted in any chromosome was that in which the most distal known gene was found. Genes shown by Green as of uncertain location (non-italicized on her map) were also discounted. If any gene locus fell exactly on one of the transecting lines that locus was always considered as being distal to the line.

When the numbers of segments containing 0, 1, 2, etc., gene loci were compared

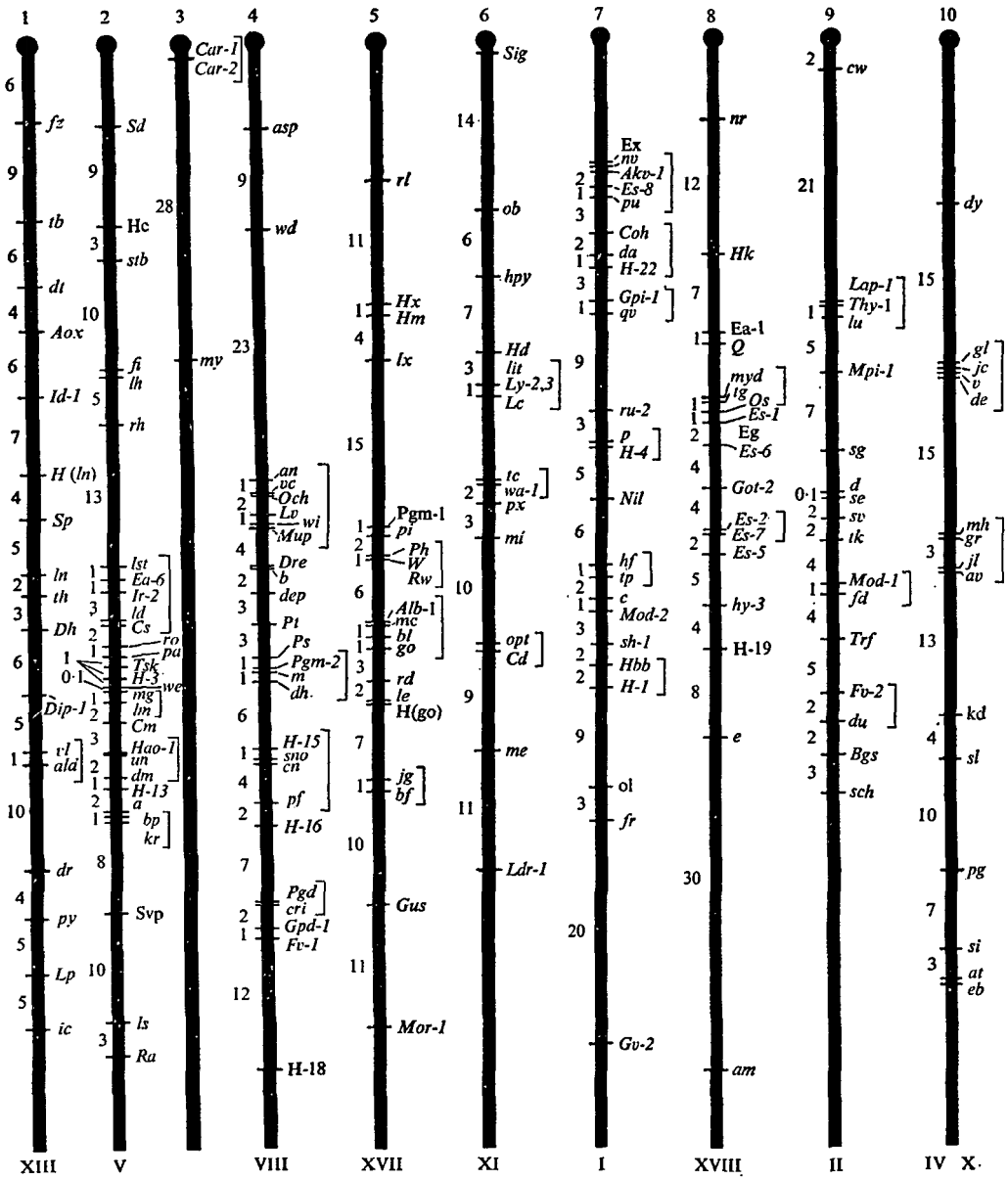

Fig. 1. For legend see facing page.

https://doi.org/10.1017/S0016672300016980 Published online by Cambridge University Press

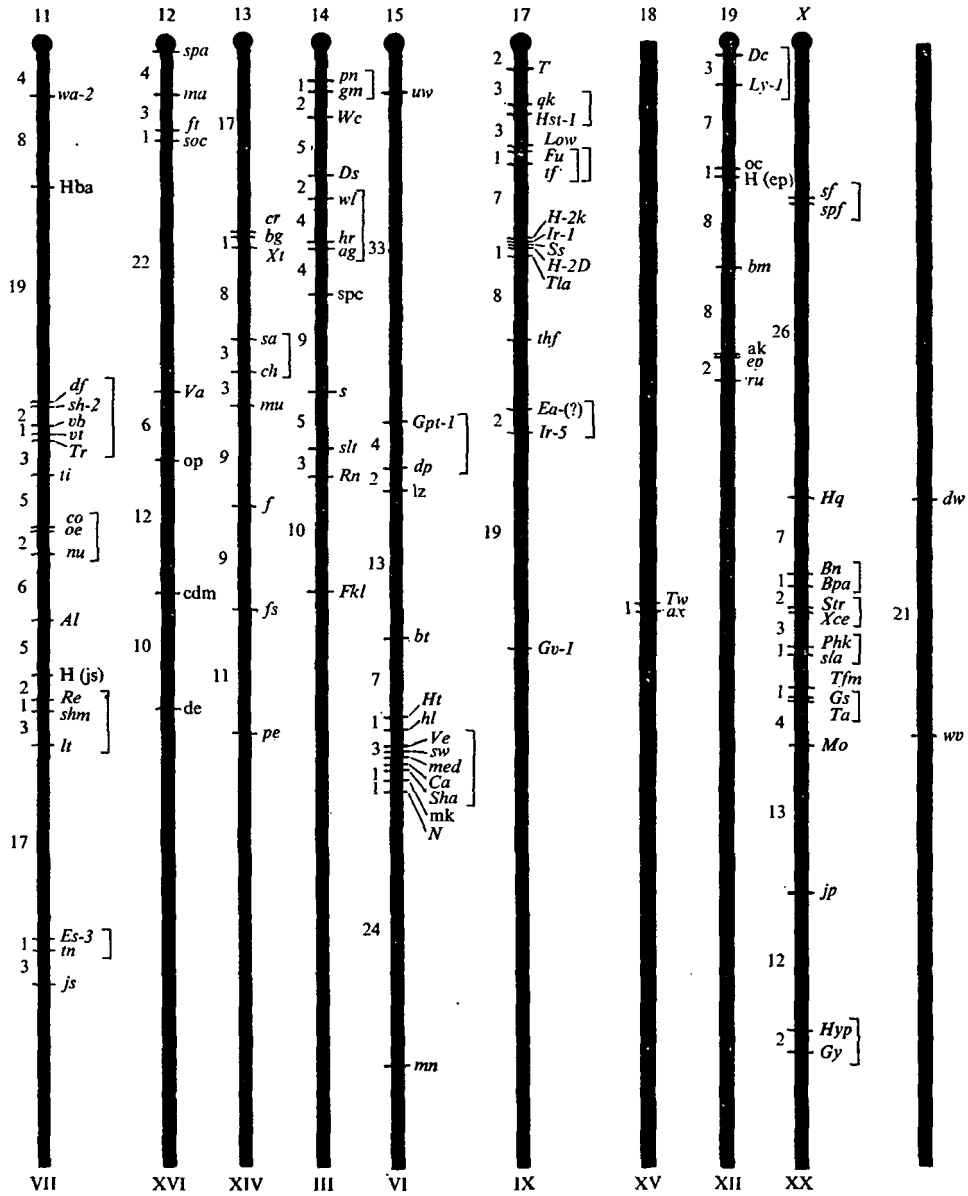

Fig. 1. Map of the known located genes of the mouse by Green (1975). Arabic numerals are chromosome numbers; Roman numerals are linkage group numbers. Loci whose order is uncertain are not italicized; brackets indicate that the order within the group is not known. Knobs indicate the centromeres. The loci of *H-18*, *Pgm-1*, *H(go)*, *Gv-2*, *Eg*, *H-19*, *kd*, *Hba*, *H(js)*, *op*, *cdm*, *de*, *spc*, *lz*, *mk*, *Gv-1*, *oc* and *H(ep)* were omitted from the data presented in Table 1 and Fig. 2 on account of uncertainty concerning their position, together with the proximal empty segments in chromosomes 2, 4, 5, 7, 8 and 10.

https://doi.org/10.1017/S0016672300016980 Published online by Cambridge University Press

with a Poisson distribution, by Fisher's index of dispersion test, the difference was clearly statistically significant (Table 1). Since the quantity $\sqrt{(2\chi^2)} - \sqrt{(2n-1)}$ is distributed as a normal deviate the probabilities are clearly very low. The deviation was due to an excess of segments in the extreme classes (0, 3, 4+), i.e. in the direction expected if there were clusters of genes on chromosomes. The clustering did not appear to be distributed evenly over all chromosomes, rather there were certain chromosomes, namely chromosomes 2, 4, 5, 6, 7, 10, 11, 15, 17 and *X*, which carried one or more clusters of at least 4 genes (Figs. 1 and 2).

Table 1. *Numbers of chromosome segments carrying 0, 1, 2, etc., gene loci*

	Segment size		Excluding outliers	
No. of loci	5 cM	10 cM	5 cM	10 cM
0	87	15	78	12
1	76	37	73	34
2	37	20	37	20
3	17	17	17	17
4	11	12	11	12
5	1	5	1	5
6	3	3	3	3
7	—	2	—	2
8	—	2	—	2
9	—	1	—	1
10	—	1	—	1
$\bar{x}$	1·16	2·33	1·20	2·44
$\Sigma(x-\bar{x})^2$	376·41	477·44	363·8	454·73
n	232	115	220	109
$\sqrt{(2\chi^2)} - \sqrt{2n-1}$	3·96	5·11	3·67	4·62
P	<0·0001	<0·0001	<0·0001	<0·0001

A possible explanation for the deviation, other than gene clustering, was an excess of observations in the O class, due to the inclusion of empty segments between the centromere and the nearest located gene. Possibly the position of the centromere might have been inaccurately estimated, or possibly there might be some reason for absence of known genes near to the centromere, e.g. the presence of centromeric heterochromatin. In fact there was only one chromosome, no. 13, in which empty segments in this position were counted, and hence this factor is unlikely to have been important.

Another possible source of error arises from certain outlying genes, which are relatively little known and difficult to work with, e.g. the genes *asp* on chr. 4, *am* on chr. 8, and *mn* on chr. 15. Because of the difficulties, such as reduced viability or difficulties in scoring, in working with these genes it is possible that, although the linkage data on which their locations are based appear sound (Collins, 1970; Meredith, 1971; Wallace & Mallyon, 1972), further work would show that in fact the disturbed segregations found were not due to genetic linkage. For the gene *Gv*-1 this has already occurred, since Stockert *et al.* (1976) have reported that the apparent association of *Gv*-1 with chr. 17 markers must be attributed to quasi-linkage rather than true linkage. The genes *Gv*-1 and *Gv*-2 have therefore been

omitted from the data presented in Table 1. The right-hand columns present the data when the outliers *asp*, *am* and *mn* are also omitted; the results still differ strongly from a Poisson distribution. In making this test we have discounted the possibility of any empty segment at the proximal or distal ends of any chromosome, and on this basis one may consider the test conservative. Thus, it seems reasonable to conclude that the distribution of the known genes on the mouse linkage map is indeed non-random.

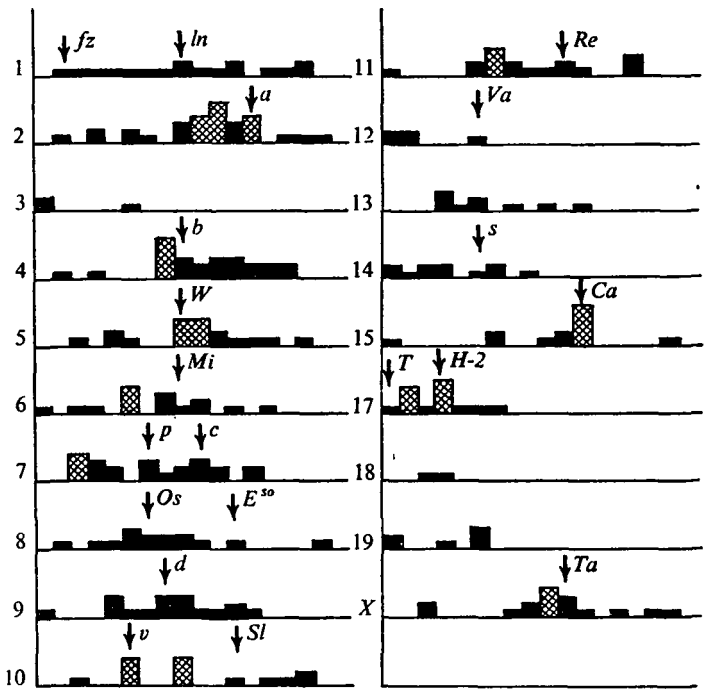

Fig. 2. Histogram of numbers of genes in 5 cM chromosome segments, in relation to common marker genes (arrows). Clusters of 4 or more genes are shown hatched.

A histogram of the number of located genes in 5 cM segments (Fig. 2) seems to indicate that, in some chromosomes (e.g. chromosome 2) segments containing 4 or more genes (i.e. clusters) tend to be located together. This was confirmed by repeating the statistical test, using a grid of lines located at 2 cm intervals (i.e. 10 cM). Again, there was a statistically highly significant departure from a Poisson distribution (Table 1), whether or not doubtful outliers were excluded.

Since there are a large number of known genes of the mouse which are not yet located, the apparent clustering might possibly be an artifact, if the locations of genes near to common, well known and widely used marker genes were found more quickly. It would then not be true to say that the known located genes were a random sample of all genes. If clustering were indeed due to preferential discovery of genes near common markers one would expect that:

(*a*) markers should lie inside clusters;

https://doi.org/10.1017/S0016672300016980 Published online by Cambridge University Press

(*b*) located genes should be distributed equally on each side of markers (providing the marker was not near one end of the chromosome);

(*c*) all clusters should contain a marker.

Genes which might be considered as 'markers' would include those which:

(*a*) occur in mutant form in certain standard inbred strains, e.g. *a*, *b*, *c*, *d*, *ln*;

(*b*) are of common distribution, according to the *Mouse News Letter*;

(*c*) are used in linkage testing stocks in major laboratories, e.g. *Ca*, *Re*, *T*.

The locations of some marker genes meeting these criteria (shown in Fig. 2) suggest that the apparent gene clustering was not due to preferential detection of genes near them, since the three expectations mentioned above are not met. First, not all markers lie inside clusters. The very well known genes dilute, *d*, and leaden, *ln*, which occur in the standard strains DBA and C57L, do not lie inside clusters (if these are defined as groups of 4 or more gene loci within 5 cM). Secondly, in those cases in which the markers lie within a cluster, the nearby genes are not always equally distributed about the marker. In particular, the agouti locus, *a*, and the loci of *Ta* and *T* are highly eccentrically situated in their clusters. Thirdly, not all clusters contain a marker. On chromosome 7 there is a group involving *nv*, *Akv-1*, *Es-8* and *pn*, on chromosome 10 a group of *mh*, *gr*, *jl*, *av*, and on chromosome 11 a group of *df*, *sh-2*, *vb*, *vt*, *Tr* (Fig. 1) and none of these genes could be described as really well known markers. Thus, it seems unlikely that the apparent clustering of genes on the mouse linkage map is an artifact due to preferential detection of genes near good markers.

Another possible artifact would arise if a particular characteristic was at first attributed to a single gene and then later, as knowledge advanced, deemed to be determined by a group of closely linked genes. An example of this is provided by the *H-2* histocompatibility system, originally considered as a single locus, but now considered as a complex of many loci (Klein, 1975) and represented by four genes on Green's map, *H-2K*, *Ir-1*, *Ss* and *H-2D*. However, *H-2* appears to be almost the only example of this kind on the map, most clusters, including those already mentioned, consisting of independently discovered genes of very diverse functions and hence it seems unlikely that any artifact arising in this way would be large enough to disturb the observed picture significantly. Even in the case of *H-2*, the *Ir-1* and *Ss* loci were discovered independently of each other and the two histocompatibility loci *H-2K* and *H-2D*, and hence this constitutes a true gene cluster.

2. DISCUSSION

Evidence has been presented that the known located mouse genes are not distributed randomly, but rather that there is a tendency to clustering of genes in certain regions of some chromosomes. This indicates that, for one reason or another, chiasmata are not randomly located on mouse chromosomes. The mean number of chiasmata per nucleus in the mouse is quite low, estimates ranging from approximately 27 to 30 in female meiosis (Ford & Evans, quoted by Evans & Phillips, 1975; Henderson & Edwards, 1968; Polani, 1972) and from 22 to 26 in males (Kyslikova & Forejt, 1972; Polani, 1972; Searle, Berry & Beechey, 1970),

with the number in any bivalent varying from 1 to 3. This implies that there is strong chiasma interference within bivalents, since once a single chiasma has formed in one of the smaller bivalents the formation of another is highly improbable. Nevertheless, random location of a single chiasma within such a chromosome could still occur, if all points along the chromosome had the same capability for forming a chiasma, and the points of initiation of chiasma formation were distributed randomly. Since the present results indicate that the distribution of chiasmata, at least in some chromosomes, is not random then one or other of these conditions is not met.

We have noted that the clustering is not obvious in all chromosomes. It is possible that the apparent differences between chromosomes are spurious, and due simply to the fact that insufficient genes have so far been located to show the clustering in all chromosomes. However, two chromosomes in which clustering is not obvious, chromosomes 1 and 9, each carry as many as 17 located genes, which would seem adequate to detect clustering if it was present. Therefore, it does seem necessary to consider the absence of clustering in some chromosomes as perhaps a real effect. One may consider whether there may be particular regions of chromosomes in which chiasmata can only form with difficulty even if no others are present. Such chromosomal differentiation, if it existed, might be associated with the differentiation into light or dark bands, after quinacrine or Giemsa staining. It is noticeable that the chromosomes with clustering (numbers 2, 4, 5, 6, 7, 10, 11, 15, 17) tended to have marked light G-bands, suggesting that light G-bands might be regions of low chiasma formation (Nesbitt & Francke, 1973). However, some other chromosomes also have marked light bands. Hence, it is not clear at present whether the chromosomal differentiation detected by G-banding has any relevance to chiasma formation.

Preferential initiation of chiasma formation at certain points might also lead to differences among chromosomes in degree of clustering, according to the chromosome length and mean number of chiasmata for a particular bivalent. Take for instance the case where chiasma formation begins at or near the chromosome ends (as suggested recently by de Boer & Groen (1974) and Polani (1972)). In a small chromosome, in which only one chiasma forms, then if initiation of chiasmata occurred equally from either end there would be relatively little departure from random distribution. If, however, initiation occurred preferentially from one end, then there would be apparent crowding of genes at the other end, where chiasmata were prevented by interference. In chromosomes where the mean number of chiasmata approaches 2·0 initiation occurring equally from both ends would lead to clustering of genes in the middle of the chromosome, and wide spacing at each end, whereas in chromosomes of such a length as to form intermediate numbers of chiasmata (1 to 2 or 2 to 3), clustering would be less obvious than in those with exact numbers as the regions of chiasma interference would depend on whether 1, 2 or 3 chiasmata occurred in any given case. Thus, the difference between chromosomes 1 and 2 might be explained in terms of number of chiasmata, if chromosome 1 had a mean of more than 2 chiasmata and hence weak clustering,

and chromosome 2 a mean near to 2 and hence strong clustering in the centre. Chromosomes *X*, 4 and 5, which also have clustering in the centre might have a mean number of chiasmata approaching 2, whereas chromosomes 15 and 17, with clustering towards one end, might have a mean of only about one chiasma. It is interesting in this context that Polani (1972) found that, in male mice in which chrs. 6 and 15 were joined in the Robertsonian translocation Rb1Ald, chr. 15 almost always had a single distal chiasma, whereas in chr. 6 there were occasionally (4 % of cases) two chiasmata and a single chiasma might be proximal or distal.

Thus, it would be possible to account for the observed clustering by various factors, such as preferential initiation of chiasmata at one or other end, or inhibition of chiasmata in certain regions and further work is needed to establish the importance of any given factor.

As a result of this gene clustering one would expect to find poor correspondence between the genetic map distances, derived from recombination, and the cytogenetic map, derived from the location of translocation and inversion break points and deletions. An example of this is seen in chromosome 17, in which the translocation break points of T(9; 17)138Ca and T(1; 17)190Ca are genetically not far removed, T190 being close to *tf* (Lyon, 1967 and unpublished) and T138 3 units distal to *H-2* (Klein & Klein 1972), so that the distance between them is about 10 units. Cytogenetically, however, they are widely apart, the T190 breakpoint being in band 17C and T138 in band 17E1 (Miller & Miller, 1972).

Another phenomenon concerning chromosome 17 and translocations T138 and T190 which may be explicable in terms of gene clustering is the enhancement of recombination between the loci of *T*, *tf* and *H-2* in the presence of these translocations (Lyon & Phillips, 1959; Klein & Klein, 1972; Lyon, 1967). In a normal chromosome the loci may lie in a region of 'clustering' or few chiasmata, and the change of pairing in a translocation heterozygote may alter chiasma localization and hence the observed recombination. Similarly, Phillips (1968, 1970) and Egorov & Blandova (1968) have noted marked changes in recombination values in chromosome 2, in the presence of the allele A^s. This allele at the *a*-locus lies in a region of strong clustering and in its presence recombination between the loci of *a* and *pa* is suppressed, whereas that between *a* and *bp* is enhanced, the order of loci being *pa -a-bp*. This could be explained if A^s tends in some way to alter the localization of chiasmata. Unusual sex differences in recombination in certain chromosome regions might also be explained by chiasma localization. In general, recombination values from male heterozygotes are lower than those from females (Dunn & Bennett, 1967), in agreement with the fewer chiasmata in male meiosis. The differences are not uniform over all chromosome intervals, however, and in some regions recombination is actually higher in males. A particular example is in chr. 15 (Dunn & Bennett, 1967; Wallace & Mallyon, 1972) in the neighbourhood of the locus of *Ca*, which we have already seen to be a region of tight clustering. Possibly, a reduction of chiasmata in other parts of this chromosome in males could partially relieve the inhibition of chiasmata in this region and so lead to

higher recombination. Polani (1972) found that, in addition to the generally higher number of chiasmata in females the detailed distributions into proximal, distal, and single or double chiasmata were different in the two sexes.

Thus, further understanding of chiasma interference and other factors affecting the localization of chiasmata on mouse chromosomes may lead to the elucidation of a range of cytogenetic phenomena.

REFERENCES

Boer, P. de & Groen, A. (1974). Fertility and meiotic behaviour of male T70H tertiary trisomics of the mouse (*Mus musculus*). A case of preferential telomeric meiotic pairing in a mammal. *Cytogenetics and Cell Genetics* **13**, 489–510.

Collins, R. L. (1970). A new genetic locus mapped from behavioral variation in mice: audiogenic seizure prone (*asp*). *Behavioural Genetics* **1**, 99–109.

Darlington, C. D. (1965). *Cytology*. London: J. and A. Churchill Ltd.

Dunn, L. C. & Bennett, D. (1967). Sex differences in recombination of linked genes in animals. *Genetical Research* **9**, 211–220.

Egorov, I. K. & Blandova, Z. K. (1968). Private communication. *Mouse News Letter* **39**, 41 & **41**, 48.

Evans, E. P. & Phillips, R. J. S. (1975). Inversion heterozygosity and the origin of XO daughters of *Bpa*/+ female mice. *Nature* **256**, 40–41.

Green, M. C. (1975). Private communication. *Mouse News Letter* **53**, 10.

Henderson, S. A. & Edwards, R. G. (1968). Chiasma frequency and maternal age in mammals. *Nature* **218**, 22–28.

John, B. & Lewis, K. R. (1965). The meiotic system. *Protoplasmatologia* **6**.

Klein, J. (1975). *Biology of the Mouse Histocompatibility-2 Complex*. Berlin: Springer-Verlag.

Klein, J. & Klein, D. (1972). Position of the translocation break T(2; 9)138Ca in linkage group IX of the mouse. *Genetical Research* **19**, 177–179.

Kyslikova, K. & Forejt, J. (1972). Chiasma frequency in three inbred strains of mice. *Folia Biologica, Praha* **18**, 216–218.

Lyon, M. F. (1967). Private communication. *Mouse News Letter*, **36**, 35.

Lyon, M. F. & Phillips, R. J. S. (1959). Crossing-over in mice heterozygous for t-alleles. *Heredity* **13**, 23–32.

Meredith, R. (1971). Private communication. *Mouse News Letter* **45**, 31.

Miller, D. A. & Miller, O. J. (1972). Chromosome mapping in the mouse. *Science* **178**, 949–955.

Nesbitt, M. N. & Francke, U. (1973). A system of nomenclature for band patterns of mouse chromosomes. *Chromosoma* **41**, 145–158.

Phillips, R. J. S. (1968). Private communication. *Mouse News Letter* **39**, 25.

Phillips, R. J. S. (1970). Private communication. *Mouse News Letter* **42**, 27.

Polani, P. E. (1972). Centromere localization at meiosis and the position of chiasmata in the male and female mouse. *Chromosoma* **36**, 343–374.

Searle, A. G., Berry, R. J. & Beechey, C. V. (1970). Cytogenetic radiosensitivity and chiasma frequency in wild-living male mice. *Mutation Research* **9**, 137–140.

Stockert, E., Boyse, E. A., Sato, H. & Itakura, K. (1976). Heredity of the G_{IX} thymocyte antigen associated with murine leukemia virus. Segregation data simulating genetic linkage (quasi-linkage). *Proceedings of the National Academy of Sciences, U.S.A.* **73**, 2077–2081.

Wallace, M. E. & Mallyon, S. A. (1972). Unusual recombination values and the mapping of the lethal miniature in the house mouse. *Genetical Research* **20**, 257–262.