PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Experimental Evidence for a New Genetic Phenomenon

Margaret E. Wallace

Phil. Trans. R. Soc. Lond. B 1958 **241**, 211-254
doi: 10.1098/rstb.1958.0003

EXPERIMENTAL EVIDENCE FOR A NEW GENETIC PHENOMENON

By MARGARET E. WALLACE

Department of Genetics, University of Cambridge

(*Communicated by Sir Ronald Fisher, F.R.S.—Received* 26 *November* 1956 *—Revised* 25 *July* 1957)

CONTENTS

During 1950 and 1951 I observed a linkage-like association between a chromosome III and three chromosome V markers in a laboratory stock of the house mouse. Towards the end of 1951 Dr Donald Michie's theory of centromere attraction was communicated to me privately. He proposed it as a possible explanation of a similar anomaly observed by several workers in hybrids from subspecific crosses, suggesting that, in such hybrids, the paternal and maternal centromeres tend to segregate to opposite poles at the first meiotic division.

It appeared reasonable to suppose that centromere differences are permanent, and that the anomaly I had observed in laboratory mice may be due to the non-random segregation of such differentiated centromeres. The term 'affinity' was coined for the new phenomenon.

Vol. 241. B. (Price 14*s*.) 27 [Published 10 April 1958

I found further cases of quasi-linkage in data from experiments designed for other purposes, and set up two series of experiments to test the hypothesis of affinity. One of these was designed to discover whether the association between markers in linkage groups III and V can be explained on this basis, and, if so, to produce data from which the position of the centromere in chromosome V may be found. The design of this investigation and the analysis of its results are given here.

The data are found to agree well with an affinity interpretation, and to disagree with other interpretations such as viability interaction, chromosomal linkage, and translocation. A map showing the position of the centromere in chromosome V in relation to its markers is constructed from these data, and this is discussed in the light of the evidence for centromere position available from multiple linkage backcross data. These two types of data agree remarkably well when their inherent defects and the simplicity of the assumptions made are taken into account; but it is still desirable to obtain further evidence that the point in linkage group V responsible for its association with linkage group III is in fact the centromere. None the less, it may be said that the experiment herein described provides very striking evidence for the existence of a new phenomenon involving non-random segregation of unlinked markers, and that affinity appears to be the best interpretation.

The theory of affinity is elaborated with special reference to the manipulation of genetic material, and to the use of affinity data of various kinds in the mapping of the centromere. An account is given of a process for obtaining metrical maps consistent both with affinity data and with data on linkage and interference, and such a map is obtained for chromosome V.

I. Introduction

Mendel's concepts of particulate segregation and of its mathematically definable nature remain fundamental to genetics after more than half a century of extensive observation and rapid theoretical advance. Only one major development of his general theory of segregation has been made, namely linkage. Other phenomena sharing some of the qualities of linkage have also been observed; in various species, of various kinds, and assignable to various causes other than the chromosomal relation to which the name 'linkage' was given. None is sufficiently widespread among living things or has sufficient impact on the general theory of genetics to be regarded as a development of the same importance as linkage itself, yet all contribute to a general picture of the structure and action of the dividing cell. Many pieces have still to be found before the jigsaw is complete and the picture wholly coherent. In so far as the importance of a new observation cannot be judged until it has found its proper place in the puzzle, it needs to be carefully described and the experimental evidence for it cautiously analysed and set out for scrutiny.

At the risk therefore of prolixity, the history of the present observation in the house mouse, and the analysis of the first experiment designed to investigate it, are given in some detail.

History

In 1950, in an attempt to find a linkage for fidget (*fi*), the mutants caracul (*Ca*, linkage group VI), dominant pied (*W*, linkage group III) and Danforth's short tail (*Sd*) were introduced to a stock of fidget mice. The first few backcross matings showed unmistakeable linkage of *Sd* and *fi* (Wallace 1950) and these, with other factors, became linkage group V. They also showed less striking but significant indication of linkage of these factors with *Ca* and *W*. The results of the first two 'tricoupling' backcrosses for *Ca*, *W* and *fi* are shown in table 1.

A balanced three-point linkage program for linkage group V using a^t (tan belly, agouti locus), *fi* and *Sd* was started (the results of which are now available, Wallace 1957). *W* was

kept segregating in as many as possible of the matings used in this program, in order to investigate the nature of its association with chromosome V. A summary was therefore maintained of the segregation of *W* with the V markers, whether the matings in which it occurred were preparation matings or linkage backcrosses from the point of view of the linkage group V program; hence data were included from matings in which *W* would be

TABLE 1. QUASI-LINKAGE OF *Ca*, *W* AND *fi*

	(*P*)		(*fi*)		(*Ca*)		(*W*)		total
	Ca *W*	*fi*	\+	*Ca* *W* *fi*	*W*	*Ca* *fi*	*Ca*	*W* *fi*	
mating 1	4	8	5	3	3	5	4	1	33
mating 2	13	17	8	6	9	8	5	12	78
	17	25	13	9	12	13	9	13	111
totals	42		22		25		22		

recombination values

Ca-*W*	*Ca*-*fi*	*fi*-*W*
42·34 %	42·34 %	39·64 %

χ^2 test for quasi-linkage: χ^2 tests fit of $(22+25+22):(42)$ to 3:1 and $= 9{\cdot}7568$ for 1 d.f. $0{\cdot}01 > p > 0{\cdot}001$. (The single-factor ratios show no significant disturbance.)

Note. In tables 1, 9, 18, 21 and 22, the segregation headings are as follows. The pair of genotypes headed '*P*' is parental, i.e. wholly non-recombinant. The pair of genotypes headed '*Ca*' '*W*' or with any other mutant symbol is the pair produced by that mode of gamete formation which exchanges, in the multiple heterozygote, the mutant indicated with its normal allele. This is the notation introduced for multi-point linkage in Wright (1947).

TABLE 2. QUASI-LINKAGE OF *W* WITH CHROMOSOME V MARKERS

	W-a^t					*W*-*fi*					*W*-*Sd*				
	W a^t	*W* *a*	\+ a^t	\+ *a*	total	*W* \+	*W* *fi*	\+ \+	\+ *fi*	total	*W* *Sd*	*W* \+	\+ *Sd*	\+ \+	total
coupling backcrosses	42	41	27	25	135	129	65	102	108	404	180	152	138	171	641
repulsion backcrosses	64	72	53	61	250	28	26	35	19	108	15	34	29	27	105
non-recombinants: recombinants	192:193					298:214					414:332				
χ^2 (1:1) for 1 d.f.	0·0026					13·7813					9·0134				
recombination fractions %	50·1299					41·7969					44·5040				

N.B. There is no significant viability disturbance in any of these bodies of data. The χ^2 value (with Yates's correction) testing contingency in the 2×2 table of the form

	non-recombinants	recombinants
coupling	*a*	*b*
repulsion	*c*	*d*

applied to each of the two-point segregations above is in each case $< 0{\cdot}1$, $p > 0{\cdot}7$. Hence, the χ^2 (1:1) above are valid tests of independence and the recombination estimates sufficiently accurate for the present purpose.

potentially linked with only one or two of the V markers, as well as with all of them. For this reason, the running totals which were maintained for each segregation are not the same at any one date, and there are, in fact, more data for *W*-*Sd* than for *W*-*fi* or *W*-a^t. The linkage of *W*-*fi* remained at about 40 % and that of *W*-*Sd* fluctuated round 45 %, while the χ^2 values rose steadily, until in September 1951 the results were as shown in table 2.

The very significant χ^2 values for *W-fi* and *W-Sd* placed the association of *W* with chromosome V beyond doubt; but the explanation was not obvious. The simplest appeared to be that linkage groups III and V concern the same chromosome. However, the recombination values obtained do not allow of the placing of *W* on chromosome V in any position consistent with current theories of interference; their relations are, in fact, non-linear (figure 1).

There seemed to be no alternative to the conclusion that linkage groups III and V refer to different chromosomes but that some point, at least in V, is responsible for the association with *W*. Such a point was thought to be situated nearest to fidget since the *W-fi* recombination value is the smallest. That this point is merely the locus of a gene affecting the expression of *W* was thought very unlikely, if only because the single-factor ratios of *fi* and *Sd* are disturbed within the non-*W* as well as in the *W* classes. Further experimental work was planned in order to discern its nature more fully.

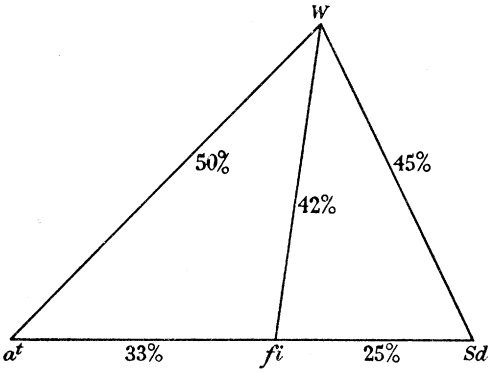


FIGURE 1. Association of *W* with a^t, *fi* and *Sd*.

At this juncture, Professor Sir Ronald Fisher received privately from Dr Donald Michie, then of Oxford, an account of a theory which, if experimentally confirmed, would explain the linkage-like association of the independent markers Maltese dilution (*d*), brown (*b*), recessive pied (*s*) and waltzer (*v*), observed by Gates (1926) in backcross progeny of a cross between two subspecies of *Mus musculus*. Michie's hypothesis was that at the first meiotic division in the hybrid, the centromeres of one subspecies tend to go to one pole and the centromeres of the other to the other pole, and that this tendency results from an attraction of centromeres of like origin, either for each other or for some polar element of the cell. Sir Ronald Fisher suggested to me that the association of *W* with chromosome V markers might be due to the same phenomenon. In this case, the point on linkage group V responsible for the *W*-V quasi-linkages would be the centromere, and the observed quasi-linkage values would be composite, being due partly to the linkage of chromosome V and III markers with their respective centromeres, and partly to an attraction between these centromeres (figure 2). Since the centromeres, in the great majority of organisms, appear to be the initial movers in the passage of the chromosomal elements from the equator to the poles of the dividing cell, it seemed reasonable to suppose that the agents both of Gates's phenomenon observed in crosses *between* subspecies and of my own observed in crosses *within* the subspecies, were in fact the centromeres.

Michie found further evidence of quasi-linkages in crosses between sub-species, in the work of Little (1927) and Green (1931), which was, on the whole, corroborative; but the theory required the proof of direct experimentation. Meanwhile I searched for and found, from experiments designed for other purposes, new quasi-linkages from crosses within laboratory stocks, that is within the subspecies. And I set up further tests. These were immediately fruitful in the production of a very significant quasi-linkage between *Sd* and *Ca*, and thus supplied experimental confirmation of the existence of the kind of quasi-linkage essential to Michie's theory. At Sir Ronald Fisher's suggestion, the term 'affinity' was adopted for both phenomena, and short simultaneous communications on the subject were made to *Nature* (Michie 1953; Wallace 1953). A brief account, including an interim summary of my *W*-V investigation and of the successful results from my crosses between stocks of laboratory mice, was given at the IXth International Congress of Genetics in 1953 (Wallace 1954*a*) and a fuller report made shortly afterwards (Wallace 1954*b*).

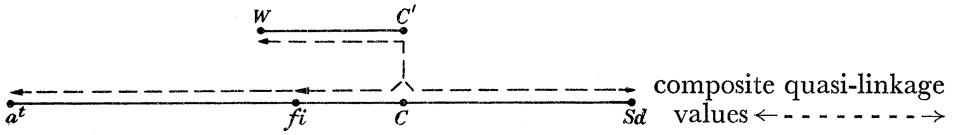


FIGURE 2. The relation of *W* with a^t, *fi*, *Sd* on the basis of 'affinity'.

It is perhaps useful in this context to stress certain points.

Michie's theory proposes that the centromeres are the causative agents for quasi-linkage in crosses between sub-species. That this is in fact so remains to be proved. Meanwhile, at his suggestion, Mr Chatterley (Oxford) obtained further hybrid material, and this has given a new case of quasi-linkage (between *Ca* and brown, *b*; private communication). This work helps to demonstrate the generality of quasi-linkage, bringing the total of subspecies involved to three, possibly four (*Mus mus domesticus*, *M.m. musculus*, *M.m. wagneri*, and Japanese waltzers of uncertain origin. The nomenclature is that of Schwarz & Schwarz 1943).

That the centromere is the causative agent in crosses *within* species also requires conclusive proof, although the experiment now to be described gives the first evidence on this point. Until such time, however, as the centromere is conclusively proved to be the causative agent in either or both cases, the term 'affinity' should not, strictly speaking, be used to describe the observed quasi-linkages, but should be reserved for Michie's theory, of which centromeric action is an integral part. Nevertheless, because this theory supplies the most obvious explanation of all the experimental results I have so far obtained, I have developed and used it in the following pages.

A second aspect of the theory needs to be discussed. Michie proposed only two kinds of centromere; further, on the data from subspecific crosses, it could not be discerned whether the differences between them are intrinsic and therefore to some degree permanent, or extrinsic, that is dependent on the cytological differences between the two subspecies involved. That there could be more than two kinds of centromere operating in one hybrid, and that the differences could be permanent were first pointed out by Sir Ronald Fisher on consideration of my work on quasi-linkages within the species. Since the germplasm of

laboratory stocks of *Mus musculus* derives from many different countries and probably from several different subspecies, there may be more than two kinds of centromere in any one stock; and also, since many generations have passed between the original introductions of different centromeres and the present stocks in which quasi-linkages have occurred, these differences must be relatively permanent. An important result of these considerations, from the point of view of manipulative genetics, is that quasi-linkages formally giving more than 50% recombination can occur (these I have defined as 'reversals' (Wallace 1953; see also II, §1)); for centromeres of like origin can arrive in the same laboratory animal from different parents, a situation which would be unusual in subspecific hybrids.

A further extension of the theory has arisen from my work on laboratory stocks. It is possible that the degree of attraction of like centromeres observable on marked chromosomes may be influenced by the number of segregating centromeres on unmarked chromosomes, that is by the degree of heterozygosity of the centromeric genotype. In laboratory stocks of mixed ancestry, animals heterozygous for the same visible markers may be of various different genotypes as regards their centromeres. On the other hand, hybrids from the same subspecific crosses may be expected to be relatively homogeneous in their centromeric genotype; hence they should give more consistent quasi-linkage values than laboratory heterozygotes. When the crosses between subspecies have been repeated, this comparison can be made; meanwhile the results of crosses between inbred lines of laboratory mice, which I hope to describe fully shortly, show sufficient heterogeneity to make the idea of a cumulative degree of attraction very plausible.

Michie's theory is expanded in his recent paper (1955); this is devoted particularly to the ideas of 'polar' and 'mutual' attraction.

II. Manipulation of genetic material

The experiments planned to investigate quasi-linkage within laboratory stocks were of two kinds. First, the program involving crosses between inbred lines of laboratory mice—hereafter called the 'outcross program.' Secondly, a program involving inbreeding of the progeny from one animal showing strong quasi-linkage between *W* and linkage group V markers—hereafter called the '*W*-V investigation'. Both were designed to demonstrate the existence of several phases of quasi-linkage, and the second, which is the subject of this paper, was designed also to provide a map showing the position, in linkage group V, of the structure responsible for quasi-linkage.

Differences between these structures are assumed to be permanent, and there is also assumed to be attraction between similar structures. The word 'centromere' will be used for them for ease of exposition. Because the experimental results so far obtained agree with the postulation of only two kinds of centromere, no more than two will at present be considered.

Before the design and analysis of the *W*-V program can be understood, some of the implications of this simple model must be explained. It is not intended at this stage to give an abstract mathematical treatment of the design and analysis of affinity experiments. The treatment given is intended merely to provide criteria for the evaluation of the evidence for affinity.

(1) *The segregation of two marked centromeres*

The conditions necessary for quasi-linkage are analogous with those for chromosomal linkage; thus, quasi-linkage will only be obtained from an individual heterozygous for the centromeres of both chromosomes—or 'doubly heterocentric'. Such individuals are of two kinds: those having received like centromeres from the same parents—whose centromeres may be termed 'in coupling' or 'convergent', and those having received unlike ones from both parents—whose centromeres may be termed 'in repulsion' or 'divergent' (table 3).

As in chromosomal linkage, divergent heterocentrics require a preparation mating, but convergent ones may be obtained direct from quasi-linkage matings of either kind; preparation mates must usually be complementary, and may be homozygous for their centromeres, i.e. 'homocentric'—or partially or fully 'heterocentric'. Finally, the degree to which like centromeres tend to pass to opposite poles, as shown by their segregation in the gametes, is a measure of their separation, the 'separation value' being analogous to the recombination value of chromosomal linkage.

TABLE 3. THE TYPES OF HETEROCENTRIC HETEROZYGOTE WITH TWO SINGLY MARKED CENTROMERES

	chromosomes III : V	phase of markers	phase of centromeres	nomenclature of heterozygote	types of segregation in the offspring
(*a*)	$\frac{A\alpha}{a\beta}\ \frac{B\alpha}{b\beta}$	coupling	convergent	CC	coupling
(*b*)	$\frac{A\alpha}{a\beta}\ \frac{b\alpha}{B\beta}$	repulsion	convergent	RC	repulsion
(*c*)	$\frac{A\alpha}{a\beta}\ \frac{b\beta}{B\alpha}$	repulsion	divergent	RD	apparent coupling (reversal)
(*d*)	$\frac{A\alpha}{a\beta}\ \frac{B\beta}{b\alpha}$	coupling	divergent	CD	apparent repulsion (reversal)
(*e*)	Partially and completely homocentric heterozygous			CH and RH	independence

Key. Symbols for markers are in Roman letters. Symbols for centromeres are in Greek letters. The convention is used here and elsewhere by analogy with that often used for chromosomal linkage, of writing centromeres from the same parents side by side, and of writing homologous chromosomes one above the other.

In contrast to the latter, however, the separation value is not an indication of the degree to which the centromere phase is maintained, for like divergent centromeres appear to attract, not to repel as do the dominant alleles in chromosomal repulsion. This means that it is not possible to forecast the segregation of the centromere markers in the offspring of a heterozygote in which only the phase of the markers is known from its parentage; its centrotype must also be known. Conversely, the full genetic constitution can only be deduced when both the parental phenotype and the segregation of the markers in the offspring are known.

A second contrast is provided by the fact that there are not two types of double heterozygote but four, for both convergent and divergent heterocentrics can receive their

dominant markers from the same or from different parents (table 3). A useful shorthand for the genetic constitutions of heterozygotes is to describe the phase of the markers first, followed by that of the centromeres (CC, RC, CD and RD, etc.). There are, however, only two types of segregation of the markers, repulsion and coupling; and it should be noted that when a heterocentric is divergent, the markers segregate as if in the opposite linkage phase from that which is specified by their distribution in the parents. This gives a quasi-linkage value formally in excess of 50 %. This situation is termed 'reversal' and is a feature of affinity which strikingly delineates it from chromosomal linkage.

A third contrast consists of the lack of phenotypic differentiation of the centrotype; for a linkage heterozygote, whether coupling or repulsion, is distinguishable by its phenotype from animals which are only singly dominant, or completely recessive, whereas in affinity heterozygotes, none of the various possible centrotypes are discernible from the phenotype; moreover, the linkage phase of a heterozygote is known from phenotypic differences between its parents, but in affinity heterozygotes the different centrotypes possible are not derivable from anything phenotypically discernible in the parents. This has several practical implications.

First, crosses between stocks known to be homocentric can produce heterozygotes of three kinds distinguishable only by their breeding performance: (i) homocentrics giving independent assortment of the markers, (ii) heterocentrics giving $<50\,\%$ recombination of the markers, and (iii) heterocentrics giving $>50\,\%$ recombination.

Secondly, the reproduction of quasi-linkage from an individual already showing it involves more than the mating of chosen phenotypes; it involves also progeny testing of individuals of similar phenotype in order to find those of a required centrotype. For example, the reproduction of a quasi-repulsion heterozygote from one which shows either coupling or repulsion requires not only the testing of suitable candidates from a preparation mating but the provision of several such matings; for not until a successfully tested candidate is found can it be certain that the preparation mating it came from is of the right centrotype.

Thirdly, in order that the expected frequencies of the various centrotypes of the offspring chosen for reproduction of the quasi-linkage may be calculated, the centrotype of the mate of the doubly heterozygous parent must be known; by analogy with chromosomal linkage, a suitable mate is not only doubly recessive but also doubly homocentric, but this feature is not discernible by its phenotype.

The reproduction of quasi-linkage is further complicated by the fact that quasi-linkage values so far observed have not diverged from 50 % by more than about 10 % in either direction. The selection by testing of a doubly heterocentric heterozygote therefore involves, ideally, the breeding from each candidate of a sufficient number of offspring to distinguish a 1:1 from a 3:2 segregation. (For a probability not more than 0·05 of being wrong, this number is 266; but in practice, decisions must often be made on lesser numbers). For this reason, it is practical to test males only.

A breeding program involving the reproduction of a quasi-linkage from the individual showing it therefore often requires the maintenance of at least three generations simultaneously, with considerable storage of individuals from matings not fully tested.

(2) *The effect and use of inbreeding*

The principle that inbreeding leads to homozygosity clearly applies to centromeres as well as to genes. In general, therefore, an inbred stock may be expected to be homocentric, and an outbred one partially or wholly heterocentric. If centromere markers are kept segregating in an otherwise inbred stock, this enforced segregation will delay the tendency to homocentricity. Once homocentricity has been reached, no quasi-linkage is possible; but if crosses between inbred stocks are made, there is the possibility—that is, if the centromeres of the stocks are different for the chromosomes concerned—that heterocentricity and therefore quasi-linkage will be regained.

Thus, a simple way to identify factors near the centromere—i.e. centromere markers—and to find evidence for affinity is to make crosses between inbred lines containing as many mutants as possible and to test the multiply heterozygous F_1 by backcrossing to a multiply recessive stock. If one (or more) quasi-linkage is found from such a source, and if the stocks are sufficiently inbred, it should be possible to make up heterozygotes capable of showing quasi-linkage with expected deviations from 50 % recombination always in the same direction whenever a cross between the two stocks is made. The outcross program now completed is designed to demonstrate this.

It may be suspected that factors showing quasi-linkage are in fact showing true linkage; this suspicion will only arise if the quasi-linkage is less than 50 %, or, when it exceeds 50 %, if it does so by an amount sufficiently small for it to be explicable in terms of interference. Confirmation that chromosomal linkage is not involved is readily obtained from the segregation of the factors within an inbred line, or from later inbreeding of the outbred stock; for in both of these cases independent assortment may be expected.

(3) *The effect of cumulative attraction*

So far, only the segregation of, and the effect of manipulative techniques on, marked centromeres has been considered. For every two chromosome-pairs involved in quasi-linkage, however, there are in the mouse eighteen other pairs of homologous chromosomes, some or all of which may be heterocentric. It is conceivable that the degree of attraction shown by two like centromeres may be influenced by the attraction also exerted on them by other like centromeres, so that the number of pairs of homologues which are heterocentric in a given heterozygote influences the degree of quasi-linkage obtained. It seems reasonable, on mechanical grounds, to expect this cumulative effect, and it should be possible to discriminate experimentally whether or not it occurs.

In general, the existence of cumulative attraction will tend to blur the clear-cut distinctions so far outlined. If two inbred stocks of mutants are used to obtain heterozygotes showing quasi-linkage, and if these stocks are not completely homocentric, then the degree of quasi-linkage obtained from different heterozygotes—even from the same mating, and certainly from different matings—may be expected to be different. During the early inbreeding of an outbred stock, there will be no sudden transition from matings showing quasi-linkage to those showing independence, but a gradual disappearance of the quasi-linkage. Thus, in any attempt to reproduce a quasi-linkage by a program which involves inbreeding, a strong divergence from 50 % obtained at the start may be expected to diminish with each successive generation.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

If there is contrast, therefore, between the relative permanence of quasi-linkage obtained immediately from outcrossed material and the relatively ephemeral quality of those obtained from a gradually inbred stock, there is reason to suspect the existence of a cumulative effect.

(4) *Mapping of centromeres—singly and doubly marked centromeres*

The observed quasi-linkage value between two markers each on different chromosomes is of little value for mapping purposes, since it is the product of three unknown parameters (the two chromosomal values between marker and centromere and the separation value between the centromeres)—although the quasi-linkage value itself does set an upper limit to the distance between either marker and its centromere.

If, however, a third marked centromere segregates simultaneously with the other two, the more definitive estimate of the relative distances of the markers from their centromeres may be made. Here there are three observable quasi-linkages compounded of three linkage and three separation values; if the latter can be assumed equal, the ratios of each map length (between marker and centromere) to the other two map lengths may be obtained. If the separation value depends on the degree of heterocentricity of the residual genotype, then this information is available *only* from simultaneous segregations; comparisons between two-point segregations are of little value. It should be noted that the assumption that separation values between centromeres from the same heterozygote are equal is one which itself requires experimental confirmation not easily obtained.

As soon, however, as two markers are available on the same chromosome, whether on the same, or on opposite sides of the centromere, and a cross is found showing quasi-linkage between these markers and a factor on another chromosome, the exact position of the centromere on the doubly marked chromosome can be found. For although there are still only three observable values, two quasi-linkages and the chromosomal value between the two linked markers, there is only one unknown separation value; and of the three unknown centromere-marker values, the upper limits for the two which concern the segments in the same chromosome (whether these segments are adjacent or overlapping) are set by the observed chromosomal value between the two linked markers. Hence, no assumption is needed about the separation value in any one mating, nor about the differences in separation values between matings due to a possible cumulative effect. If the two observed quasi-linkage values vary from one triple heterozygote to another, as they will if there is a cumulative effect, this variation does not introduce a bias into the estimation of the position of the centromere on the doubly marked chromosome. It should perhaps be pointed out that the estimation of centromere position from such material is, in general, a less complicated task than the estimation of map distances in multi-point linkage, if it can be assumed—and there is strong evidence in favour of this assumption—that there is no interference across the centromeres. It is also worth noting that in general the observed quasi-linkage and chromosomal values will not allow of a linear arrangement of the markers, so that if such a non-linear relationship is observed, this is evidence in favour of affinity and against multiple linkage. The calculation of the position of the centromere in linkage group V will be given in detail, and should serve to illustrate these points.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The most informative type of affinity mating is one in which at least two markers on each of two chromosomes show quasi-linkage; for here there are six observable values (four quasi-linkages and two chromosomal), so that all five parameters may be estimated (the four chromosomal values marker-centromere and the separation value). Thus, the position of the centromeres relative to their markers on both chromosomes may be found. Also the inter-relationship of all the quasi-linkage values cannot, under any circumstances, be linear. Such crosses, then, may provide not only valuable data for mapping purposes but also critical evidence for affinity. Material is now being built up in the hope of producing such crosses.

A further aspect of the use of work with affinity in the construction of genetical maps may be mentioned here. If it can be assumed that no chromosome has more than one centromere —and there is no evidence to the contrary at least in higher organisms—any two markers which show quasi-linkage with each other as opposed to chromosomal linkage, are on different chromosomes. Further, two markers which show strong quasi-linkage with a third marker are probably close to their centromeres, and if they assort independently of each other, they cannot be so far apart as to show 50% recombination, but must be on separate chromosomes. Affinity crosses, therefore, can be useful in confirming or breaking up existing loose or doubtful linkage groups, and in deciding whether those groups which are small enough to be on the same chromosome and too far apart to show linkage, are in fact parts of the same or of different groups.

(5) *The features distinguishing chromosomal linkage from affinity*

The purpose of the experiments designed to demonstrate affinity is to produce results which will distinguish it primarily from what it most resembles, which is chromosomal linkage, and secondarily from any other known genetic phenomenon which it may, to some extent, resemble. The characteristics distinguishing it from chromosomal linkage may be summarized as follows.

(i) recombination greater than 50%, especially if too great to be explained in terms of interference;

(ii) heterogeneity between linkage matings, due to there being matings giving 50% as well as greater or less than 50%;

(iii) disappearance of quasi-linkage on inbreeding;

(iv) non-linear relation of quasi-linkages.

Some of these characteristics may be shared by one or more known genetic phenomena such as translocation, inversion and interaction. The outcross program and the *W*-V investigation have between them produced all four, and have not produced any features inexplicable in terms of affinity.

III. The design of the *W*-V investigation

(1) *Outline*

No single mouse which contributed to the very significant quasi-linkages between *W* and the three linkage group V markers a^t, *fi* and *Sd* (table 2), and was still alive, had a large enough progeny to show, alone, a significant quasi-linkage. It was accordingly proposed

to isolate, by large-scale testing, one such animal from among these and their progeny, and to produce from him heterozygotes showing independence and heterozygotes showing all the possible kinds of quasi-linkage. The latter heterozygotes are doubly heterocentric—the CC, RC, CD, and RD types (table 3). The former are singly or doubly homocentric (CH and RH) and are thus of several different kinds; they constitute the most easily produced class, from which the less easily obtained heterocentrics must be isolated.

(2) *Material*

It was realized that this, the first genetic material available for such a program, was not ideally suitable for it, and that there would be many difficulties. First, in order to obtain sufficient numbers for deviations from 50% recombination (in either direction) to be distinguished significantly from independence, preference was given to males for testing rather than to females, and hence nearly all tester animals (multiple recessives) had to be fidget females. These animals are normally very poor performers in unselected material and tend to die after only a few litters (Grüneberg 1943–52). I had produced, by selection, a stock with relatively high fertility and viability for the three-point linkage program, but even here not more than an average of thirteen fully classified progeny could be expected per female. Each tested male was therefore given between five and ten females for most of his life, his mates being replaced as soon as they died or became infertile. In this way it was hoped that at least 100 progeny could be bred per male; and indeed the best progeny numbers were over 300.

A second difficulty was that the program involved strict inbreeding of the progeny of the selected animal, because, there being no way of knowing with certainty the centrotype of any animals but the one tested, and probably of his mates, this was the only way of predicting the path of his centromeres with any confidence; such inbreeding was expected to reduce the initial high fertility and viability of both the fidget mates and their progeny. Both these expectations were in fact realized, and a third difficulty was added; for an epidemic of mouse catarrh broke out. This catastrophe affected both the matings producing animals to be tested, and the animals testing them. The numbers of heterozygotes tested is therefore not as great as could be wished, and the progenies not nearly as large as could be produced under better conditions. With larger progenies it would be possible to obtain more clear-cut distinctions between quasi-linkage and independence than has been possible here, a situation which should be realized in future affinity work.

(3) *Provisional mapping of the centromere in linkage group V*

The method by which the various types of heterozygote were to be produced was decided upon in this way: from the quasi-linkage values produced by the selected animal, and from the linkage values so far obtained in the three-point backcross, a tentative map was made showing the chromosome V centromere in relation to its markers. This in turn provided provisional linkage and quasi-linkage values for the elaboration of the gametic output of this animal and of his progeny, and these values were used to find a simple and quick way of breeding the required types of heterozygote. Quasi-linkage data from all animals of this program were pooled and calculations based on them provide the maps given at the end of the paper.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

To map the centromere on a doubly marked chromosome, where only one marker on another chromosome is available, the following argument may be used.

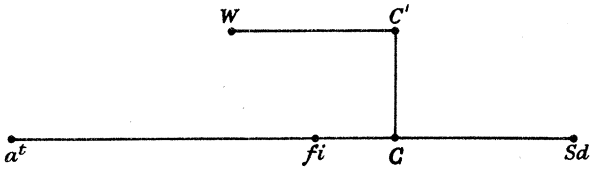
If AB and BC are adjacent segments and there is assumed to be no interference between them, the relation between the three recombination fractions pertaining to them may be written

$$\text{A-C} = \text{B-C}\ (1-\text{A-B}) + \text{A-B}\ (1-\text{B-C}),$$

where A-B, A-C and B-C represent the recombination fractions in segments AB, AC, and BC, respectively. This relation, first proposed by Trow (1913) may be expressed more usefully for the present purpose as

$$(1-2\text{A-C}) = (1-2\text{A-B})\ (1-2\text{B-C}). \tag{1}$$

In the *W*-V association we have a conceptually linear system of six points



and if there is assumed to be no interference across the centromere, three equations of the form above (1) are available

$$(1-2W\text{-}Sd) = (1-2W\text{-}C')\ (1-2C'\text{-}C)\ (1-2Sd\text{-}C), \tag{2}$$

$$(1-2W\text{-}fi) = (1-2W\text{-}C')\ (1-2C'\text{-}C)\ (1-2fi\text{-}C), \tag{3}$$

$$(1-2W\text{-}a^t) = (1-2W\text{-}C')\ (1-2C'\text{-}C)\ (1-2a^t\text{-}C). \tag{4}$$

The terms on the left are calculated directly from the three observable quasi-linkage values. Each concerns a particular V marker and is proportional to a corresponding term on the right concerning the unknown chromosomal linkage value between the V marker and the centromere. This being so, these unknown values can be found with the further aid of the observed chromosomal values between the V markers, as follows:

(*a*) *Ignoring relations with* a^t

The relation for chromosome V corresponding to (1) above is

$$(1-2fi\text{-}Sd) = (1-2fi\text{-}C)\ (1-2Sd\text{-}C). \tag{5}$$

From equations (2), (3) and (5), the following identities are available:

$$(1-2fi\text{-}C)^2 = (1-2W\text{-}fi)\ (1-2fi\text{-}Sd)/(1-2W\text{-}Sd), \tag{6}$$

$$(1-2Sd\text{-}C)^2 = (1-2W\text{-}Sd)\ (1-2fi\text{-}Sd)/(1-2W\text{-}fi), \tag{7}$$

$$(1-2W\text{-}C')^2\ (1-2C\text{-}C')^2 = (1-2W\text{-}fi)\ (1-2W\text{-}Sd)/(1-2fi\text{-}Sd). \tag{8}$$

Thus it can be seen that the recombination fractions *fi*-*C* and *Sd*-*C* can be obtained without knowledge of the separate values *W*-*C'* and *C*-*C'*; and also that the average value of the compound *W*-*C'*-*C* is calculable. Although the ingredients of this compound cannot be determined, an upper limit, namely the whole value of the compound, may be set to *C'*-*C*. Further, if *C'*-*C* should vary from mating to mating, this fact does not disturb the values *fi*-*C* and *Sd*-*C* which are estimated without using it. Hence, all matings showing

quasi-linkage can be pooled for the estimation of centromere position in V, even though *C′-C* may be different for each of them. Such differences may arise from differences in the residual genotype, as has been pointed out. They can also be conceived as arising from differences in the separation value as between heterozygotes with convergent centromeres and those with divergent ones, although there is no obvious reason why this should be the case. It is clear that, even if it is the case, it is legitimate to pool data from all kinds of segregations; those with recombination values exceeding 50 % $(1-y)$ can be expressed as the corresponding values less than 50 % (y).

For the present paper, quasi-linkage and linkage relations with a^t will be ignored for mapping purposes, except as a check upon the position of centromere V as calculated without it. This is because the observed quasi-linkages with a^t are so near 50 %, that enormous error would be involved in any calculations using it. But for the case where this is not so, it is worth noting the appropriate mapping procedure.

(*b*) *Including relations with* a^t

Two further equations of the form of (5) are available, making six equations in all for the estimation of four unknown parameters. There are thus two degrees of freedom available for testing goodness of fit. A maximum-likelihood method of estimation can be used, but a simple procedure is as follows: a trial value can be used for each of the three last terms in equations (2), (3) and (4); these must be based on a guessed position of centromere V, and this in turn must be consistent with interference relations within the segments *a^t-fi* and *fi-Sd*, for which the observed linkage values *a^t-fi*, *fi-Sd* and *a^t-Sd* are available. If the fit is not good, a second and further values may be tried until the constant, $(1-2W\text{-}C')(1-2C'\text{-}C)$ is, as nearly as may be desired, the same for each of the three equations.

(4) *Mapping with two doubly marked chromosomes*

It may be of interest to note here the appropriate relations for the case where both chromosomes have at least two markers. Material involving several chromosomes is being built up for this purpose: one stock, with *W*, *fi* and *Sd*, includes luxate (*lx*) as the second chromosome III marker. With these four points, four equations of the form (2) and (3) are available, and these provide the relation

$$\frac{(1-2lx\text{-}fi)}{(1-2lx\text{-}Sd)} = \frac{(1-2W\text{-}fi)}{(1-2W\text{-}Sd)}. \tag{9}$$

This relation is a crucial test of affinity, because it holds only if the 'centromeres' are fixed points and if there is no interference across them; therefore if it does hold, this is evidence that the points involved *are* the centromeres.

These four equations together with a further equation of the form of (5) now available with *lx*, and with equation (5) itself, provide the relation

$$(1-2C'\text{-}C)^4 = \frac{(1-2lx\text{-}fi)\,(1-2lx\text{-}Sd)\,(1-2W\text{-}fi)\,(1-2W\text{-}Sd)}{(1-2lx\text{-}W)^2\,(1-2fi\text{-}Sd)^2}. \tag{10}$$

This determines *C′-C*, whose value, used in appropriate equations of the form (8), allows the centromeres in both chromosomes to be mapped in relation to their markers. Alternatively, equations (6) and (7) may be used with two further equations of this form now available.

(5) *Gametic frequencies from heterocentric heterozygotes*

The animal selected as progenitor of the various types of heterocentric heterozygote was given the letter 'H'. He segregated in a^t, *fi* and *Sd* as well as *W*. It was not thought practicable to test only those of his descendants which segregated in *all* these factors; for if such a limit had been imposed, there would probably have been insufficient numbers for testing. Such a procedure, if practicable, would have produced more complete data for mapping, but, as has been pointed out, the program was not designed primarily for this purpose but as a test of the existence of affinity.

Since *W-fi* had been consistently the closest quasi-linkage (table 2), it was decided to demonstrate the various types of heterocentricity with these two loci at least, and gametic frequencies are calculated as if only these two were segregating. In practice, *Sd* was kept segregating as often as possible with *fi* and occasionally in place of it, and a^t was included only when available. That there was, however, no bias in favour of *W-fi* quasi-linkages can be seen from table 10; actually *Sd* occurs more often in animals not segregating for both factors than does *fi*.

To obtain the modes of gamete formation, the four points may be regarded as linear: *W-C′-C-fi*. Then, as in four-point linkages, there are three simple recombination values, eight heterozygotes and eight modes of gamete formation. The latter fall into two classes, those involving an odd number of crossovers and those involving an even number ('odds' and 'evens'). The 'evens', by analogy with four-point linkage, give non-recombination for the end markers *fi* and *W*, and the 'odds' give recombination.

Three quasi-linkage values, *W*-a^t, *W-fi* and *W-Sd*, were obtained from 'H''s performance after he had produced ninety-one progeny (table 8), and the three chromosomal values a^t-*fi*, *fi-Sd* and a^t-*Sd*, from the segregations from males in the three-point program then almost finished. These values together gave a provisional map, from which the following approximations were taken:

	simple recombination or separation values, y	$1-y$
W-C′	$\frac{1}{8}$	$\frac{7}{8}$
C-C′	$\frac{1}{3}$	$\frac{2}{3}$
fi-C	$\frac{1}{10}$	$\frac{9}{10}$

These give the frequencies of the modes of gamete formation in table 4, where they are tabulated as adding to 240. The subtotal for 'evens' is 144, and that for 'odds' 96 (column 3), the ratio 144:96 being effectively the ratio $(1-W\text{-}fi):(W\text{-}fi)$ observed in 'H''s 91 progeny.

As in a chromosomal-linkage program, all (or nearly all) mates from whom maintenance of the program is intended are chosen to be complementary; because the centrotype is not phenotypically differentiated, it is seldom possible to choose complementary centrotypes, but they are the most useful to employ when choice is possible. The two forms of linkage are also found to share the same principle as regards their frequencies, namely, that the members of a fully complementary pair are produced from a complementary mating by the same mode and therefore with the same frequency. In making up a table of gametic frequencies it is therefore necessary to specify the constitution and frequency of only one member of a pair (table 5). Hence also, the frequency with which a fully complementary pair will occur as mates, when mating is at random within complementary phenotypes, is

found by squaring the frequency of the mode by which it is produced (table 4, column 5. Column 4 of table 4 expresses the ratios in column 3 as percentages, and these, squared, provide column 5)

TABLE 4. FREQUENCIES OF THE MODES OF GAMETE FORMATION FROM ANIMALS DOUBLY HETEROCENTRIC AND HETEROZYGOUS FOR *W* AND *fi*

Set (*a*): *W-fi* not recombined

1	2	3	4	5
even nos. of recombinations	frequencies	ratios	percentages	frequencies of complementary matings within set (*a*)
no recombination	$7\times2\times9$	126	87·5000	76·5625
W-C′, *C′-C*	$1\times1\times9$	9	6·2500	0·3906
fi-C, *C′-C*	$7\times1\times1$	7	4·8611	0·2363
W-C′, *fi-C*	$1\times2\times1$	2	1·3889	0·0193
		144	100·0000	

Set (*b*): *W-fi* recombined

odd nos. of recombinations	frequencies	ratios	percentages	frequencies of completementary matings within set (*b*)
C′-C	$7\times1\times9$	63	65·6250	43·0664
W-C′	$1\times2\times9$	18	18·7500	3·5156
fi-C	$7\times2\times1$	14	14·5833	2·1267
W-C′, *fi-C*	$1\times1\times1$	1	1·0417	0·0109
C′-C		96	100·0000	

grand total 240

To specify fully the frequencies of the eight complementary pairs as produced by each of the eight kinds of heterozygote, the resemblance to four-point chromosomal linkage may be considered further. In the latter system there are also eight kinds of segregation, but in the present example of quasi-linkage there are only four, this is because the heterozygotes may be arranged in pairs, the members of each pair producing the eight complementaries with identical frequencies: tables 5*A* (*a*) and (*b*) and 5*B* (*a*) and (*b*). This difference between the two systems is due to the assumption that like centromeres tend to attract and not to repel whether they come from the same or from different parents; whereas in four-point linkage the two alleles at each locus can be either in coupling or in repulsion with those at any other locus. The same limitation operates when there are more than two chromosomes, and when these are singly or multiply marked.

Thus, in chromosomal linkage there are 2^{l-1} kinds of heterozygote (where l is the number of loci), and this expression suffices for the number of modes of gamete formation and the number of kinds of segregation. But in the present model of quasi-linkage, there are $2^{(l+n)-1}$ kinds of heterozygotes and modes (where n is the number of chromosomes involved), and $2^{(l+n)-1}\div2^{n-1}$, or 2^l kinds of segregation.

Each complementary pair in table 5 is indicated by the member containing *W*. The constitution of the two chromosomes received from the heterozygous parent only is given, since to continue the program, successfully tested heterozygotes were planned to be mated with $+^{W}\beta\, fi\beta/+^{W}\beta\, fi\beta$ animals (or β stock) which could therefore supply only $+\beta\, fi\beta$ chromosomes. The numerals (i), (ii), etc., label identical chromosomal genotypes in the four tables.

It should be pointed out that the heterocentrics and the chromosomal contributions in table 5 $A(a)$ become those in $A(b)$ by the interchange of α with β; they become those in $B(a)$ by the interchange of α with β and $+^{fi}$ with fi; and they become those in $B(b)$ by the interchange only of $+^{fi}$ with fi; this exhausts all the permutations.

TABLE 5. GAMETIC OUTPUT OF ALL TYPES OF ANIMAL DOUBLY HETEROCENTRIC AND HETEROZYGOUS FOR *W* AND *fi*

recombination ('evens')	Members of complementary pairs including *W*	ratios	recombination ('odds')	Members of complementary pairs including *W*	ratios
(*A*) Those giving coupling data for *W* and *fi*: (CC and RD heterozygotes)					
		(*a*) $\frac{W\alpha}{+\beta}\frac{+\alpha}{fi\beta}$ and $\frac{W\alpha}{+\beta}\frac{fi\beta}{+\alpha}$			
no recombination	$W\alpha\ +\alpha$ (i)	126	*C'-C*	$W\alpha\ fi\beta$ (v)	63
W-C', *C'-C*	$W\beta\ +\alpha$ (ii)	9	*W-C'*	$W\beta\ fi\beta$ (vi)	18
fi-C, *C'-C*	$W\alpha\ +\beta$ (iii)	7	*fi-C*	$W\alpha\ fi\alpha$ (vii)	14
W-C', *fi-C*	$W\beta\ +\beta$ (iv)	2	*W-C'*, *fi-C*, *C'-C*	$W\beta\ fi\alpha$ (viii)	1
		(*b*) $\frac{W\beta}{+\alpha}\frac{+\beta}{fi\alpha}$ and $\frac{W\beta}{+\alpha}\frac{fi\alpha}{+\beta}$			
no recombination	$W\beta\ +\beta$ (iv)	126	*C'-C*	$W\beta\ fi\alpha$ (viii)	63
W-C', *C'-C*	$W\alpha\ +\beta$ (iii)	9	*W-C'*	$W\alpha\ fi\alpha$ (vii)	18
fi-C, *C'-C*	$W\beta\ +\alpha$ (ii)	7	*fi-C*	$W\beta\ fi\beta$ (vi)	14
W-C', *fi-C*	$W\alpha\ +\alpha$ (i)	2	*W-C'*, *fi-C*, *C'-C*	$W\alpha\ fi\beta$ (v)	1
(*B*) Those giving repulsion data for *W* and *fi*: (RC and CD heterozygotes)					
		(*a*) $\frac{W\beta}{+\alpha}\frac{fi\beta}{+\alpha}$ and $\frac{W\beta}{+\alpha}\frac{+\alpha}{fi\beta}$			
no recombination	$W\beta\ fi\beta$ (vi)	126	*C'-C*	$W\beta\ +\alpha$ (ii)	63
W-C', *C'-C*	$W\alpha\ fi\beta$ (v)	9	*W-C'*	$W\alpha\ +\alpha$ (i)	18
fi-C, *C'-C*	$W\beta\ fi\alpha$ (viii)	7	*fi-C*	$W\beta\ +\beta$ (iv)	14
W-C', *fi-C*	$W\alpha\ fi\alpha$ (vii)	2	*W-C'*, *fi-C*, *C'-C*	$W\alpha\ +\beta$ (iii)	1
		(*b*) $\frac{W\alpha}{+\beta}\frac{fi\alpha}{+\beta}$ and $\frac{W\alpha}{+\beta}\frac{+\beta}{fi\alpha}$			
no recombination	$W\alpha\ fi\alpha$ (vii)	126	*C'-C*	$W\alpha\ +\beta$ (iii)	63
W-C', *C'-C*	$W\beta\ fi\alpha$ (viii)	9	*W-C'*	$W\beta\ +\beta$ (iv)	18
fi-C, *C'-C*	$W\alpha\ fi\beta$ (v)	7	*fi-C*	$W\alpha\ +\alpha$ (i)	14
W-C', *fi-C*	$W\beta\ fi\beta$ (vi)	2	*W-C'*, *fi-C*, *C'-C*	$W\beta\ +\alpha$ (ii)	1

The assumption that the β stock was constitutionally $\frac{+\beta}{+\beta}\frac{fi\beta}{fi\beta}$ rests on the following argument. *W* and *fi* came into the stock from different sources and the first quasi-linkage matings gave coupling segregations; this suggests that the heterozygous mates were $\frac{W\alpha}{+\beta}\frac{+\alpha}{fi\beta}$, that is that the *W* stock was of one centrotype for both the III and V chromosomes and the *fi* was of another. These are conveniently labelled α and β. Although later fidgets were derived from these matings and not from the original stock, there was reason, because of the linkages and affinity relations, to expect that these were mainly β, and selection for the 'β' stock was made from those least contaminated with the 'α' stock.

'H' distinguished himself from the other tested males by giving $>50\,\%$ coupling; that is, he was a divergent heterocentric and his constitution was therefore either

$$\text{(i)}\ \frac{W\alpha}{+\beta}\,\frac{+\beta}{fi\alpha} \quad \text{or} \quad \text{(ii)}\ \frac{W\beta}{+\alpha}\,\frac{+\alpha}{fi\beta}.$$

There is no way of deciding, from his ancestry, which is the correct interpretation. However, as will be seen later, it does not make much difference to the planning of the program concerning his progeny (figure 3). Since *fi*-*C* is, however, assumed to be smaller than *W*-*C*′, we shall at present assume that *W*-*C*′ rather than *fi*-*C* crossovers took place somewhere in his ancestry and that (ii) is his constitution.

(6) *The planned pedigree of* 'H''*s descendants*

The program designed to demonstrate the various kinds of heterozygotes among 'H''s progeny was constructed as follows: (Table 5 is used in conjunction with table 4).

(*a*) CC *heterozygotes*

From table $5B(a)$, it can be seen that CC heterozygotes of only one centrotype, namely

$$\frac{W\alpha}{+\beta}\,\frac{+\alpha}{fi\beta},$$

are available in 'H's' F_1—chromosome genotype (i). The other three types of coupling heterozygote are singly or doubly homocentric and therefore give $50\,\%$ recombination. It is thus impossible to obtain more than $50\,\%$ recombination from the F_1, an expectation which, if fulfilled, is a good check on the hypothesis of affinity. The frequency of mode *W*-*C*′, which produces CC, among all modes giving F_1 heterozygotes, is $18{\cdot}75\,\%$ (table $4(b)$).

The other three heterozygotes which it was planned to demonstrate have their markers in repulsion or their centromeres divergent, or both, and therefore a preparation mating is required to produce them from 'H'.

(*b*) RC *heterozygotes*

These are either

$$\text{(i)}\ \frac{W\alpha}{+\beta}\,\frac{fi\alpha}{+\beta} \quad \text{or} \quad \text{(ii)}\ \frac{W\beta}{+\alpha}\,\frac{fi\beta}{+\alpha}.$$

The left side of table $5B(a)$ supplies the frequencies with which the complementaries, which form the preparation-matings for these two heterozygotes, are produced from 'H'. These are pairs (vi) and (vii). Pair (vii) has so low a frequency among all preparation matings, $1{\cdot}39\,\%$ (see table $4(a)$) that it may be ignored. The frequency of the other mating $\frac{W\beta}{+\beta}\,\frac{fi\beta}{fi\beta}\times\frac{+\alpha}{+\beta}\,\frac{+\alpha}{fi\beta}$ among all preparation matings, is $76{\cdot}5625\,\%$ (see table $4(a)$). One of the mates here, however, offers a slight problem; for the phenotypically normal is required to give a $+\alpha\ +\alpha$ chromosomal contribution; the non-*fi* chromosome will be α if no *fi*-*C* crossover occurs, and this will assort with a (III) $+\alpha$ chromosome if no *C*′-*C* crossover occurs; therefore $\frac{2}{9}\times\frac{9}{10}=60\,\%$ of the heterozygotes from this mating will be heterocentric.

It was at first intended to adopt the following policy for the whole program: to identify a suitable preparation mating by the performance of one of its heterozygotes; and then to

proliferate heterozygotes from the mating thus isolated so as to demonstrate that it gave a certain ratio of heterozygotes giving the quasi-linkage required to those not giving it. But the shortage of material due to catarrh made this impossible, so that only about one heterozygote per preparation mating of any kind was tested.

To find the frequency of RC heterozygotes in the F_2 it is, therefore, necessary merely to multiply the frequency of suitable F_1 matings by the frequency of RC heterozygotes from them, i.e. $76{\cdot}5625\,\% \times 60\,\% = 45{\cdot}9375\,\%$.

It should be noted that this is a slightly lower figure than would be obtained if all possibilities were taken into account (table 6*B*), so that 50 % is a handy and not too remote approximation. It should also be noted that in all cases where two generations are used to find a particular heterozygote, there is an opportunity for the centromeres to reassort in the intervening generation; thus it is possible here to find a few divergent heterozygotes in the F_2. Hence, it is *only* in the production of CC heterozygotes that the absence of quasi-linkages exceeding 50 % is a check on the hypothesis of affinity; and thus it is also advisable to check that the frequency of heterozygotes giving quasi-linkages in the opposite direction to that required, is not greater than the frequency of heterozygotes giving the required quasi-linkage.

To obtain the exact frequency of heterocentric heterozygotes in a two-generation program is a rather laborious process, of which only one example is necessary here. It requires the tabulation of the frequencies of all possible preparation matings, and the calculation of the frequencies with which each one capable of giving a heterocentric of either kind can do so. Thus, table 6*A* demonstrates that seven different genotypes of matings can give an RC heterozygote, the frequency of all such matings being actually 87·6737 % (table 6*B*). However, the frequency of each severally, except the one considered in detail above, is less than 5 %, and the calculation of the frequency of heterocentrics from each reduces further the total expectation of heterocentrics in F_2, so that the figure derived from the complete calculation, 47·7679 %, is not far from the rough approximation of 50 %.

Table 6*A* also shows that seven matings can give an RD heterozygote, the total frequency of such matings being 11·7188 % (table 6*C*). Of these only two have a frequency exceeding 4 %, and it can be seen from their constitutions that only the one with a frequency of 5·4688 % has an appreciable expectation of producing an RD heterocentric F_1; the *W* mate can produce a $W\alpha\, fi\beta$ contribution in the absence of a *W*-*C*′ crossover, the non-fidget gamete of the other mate will be $+\alpha$ only if there is no *fi*-*C*′ crossover, and this will assort with a $+\beta$ chromosome only if there is a *C*′-*C* crossover. That is, the expectation that heterozygotes from this mating will be divergent heterocentrics is $\frac{7}{8} \times \frac{1}{3} \times \frac{9}{10}$ or 26·25 %. Thus, the total expectation of RD heterozygotes in the F_2 is little more than $5{\cdot}4688\,\% \times 26{\cdot}25\,\%$ or about $1\frac{1}{2}$ %. The full calculation (table 6*C*) produces 2·0054 % and thus raises this rough approximation only slightly.

(*c*) RD *heterozygotes*

As has been noted, these are either (i) $\frac{W\alpha\, fi\beta}{+\beta\ +\alpha}$, or (ii) $\frac{W\beta\, fi\alpha}{+\alpha\ +\beta}$. Here, advantage may be taken of the fact that a previously proven centrotype, namely, the CC heterozygote isolated in the F_1, may be used as a stage in obtaining (i)—via a further preparation mating.

The most suitable mate would be a 'β' fidget, and it can be shown, from a table of calculations similar to table 6, that the total frequency of F_3RD heterozygotes from F_1CC heterozygotes with these mates would be about 40 %. It can similarly be shown that only about 5 % of the heterozygotes tested for being RD would be RC.

Unfortunately, it was found impossible to give many β females to the F_1CC heterozygotes, since the epidemic of catarrh reduced the supply so much that most of them had to

Table 6

(*A*) Frequencies of centrotypes of preparation-matings available in 'H''s F_1

	Phenotypically normal mates				
phenotypically *W-fi* mates	$\frac{+\alpha\ +\alpha}{+\beta\ fi\beta}$	$\frac{+\beta\ +\alpha}{+\beta\ fi\beta}$	$\frac{+\alpha\ +\beta}{+\beta\ fi\beta}$	$\frac{+\beta\ +\beta}{+\beta\ fi\beta}$	frequencies
$\frac{W\beta\ fi\beta}{+\beta\ fi\beta}$	76·5625 C	5·4688 —	4·2535 —	1·2153 —	87·5000
$\frac{W\alpha\ fi\beta}{+\beta\ fi\beta}$	5·4688 D C	0·3906 D	0·3038 —	0·0868 —	6·2500
$\frac{W\beta\ fi\alpha}{+\beta\ fi\beta}$	4·2535 D C	0·3038 —	0·2363 D	0·0675 —	4·8611
$\frac{W\alpha\ fi\alpha}{+\beta\ fi\beta}$	1·2153 D C	0·0868 D C	0·0675 D C	0·0193 C	1·3889
frequencies	87·5000	6·2500	4·8611	1·3889	100

Key: C, indicates that the mating gives convergent heterozygotes as well as single or double homocentrics. D, indicates that the mating gives divergent heterozygotes as well as single or double homocentrics. —, indicates that the mating gives only single or double homocentrics.

(*B*) Frequency of convergent heterocentrics among repulsion heterozygotes from matings in (*A*)

(*a*) frequencies of matings giving convergent hets (%)	(*b*) frequencies with which heterozygotes from (*a*) are convergent							(*a*) × (*b*)
	W-fi mate				normal mate			
	W-C′	*C′-C*	*fi-C*		*W-C′*	*C′-C*	*fi-C*	(%)	(%)
76·5625	1	1	1	×	1	$\frac{2}{3}$	$\frac{9}{10}$	= 60·0000	45·9375
5·4688	$\frac{1}{8}$	1	1	×	1	$\frac{2}{3}$	$\frac{9}{10}$	= 7·5000	0·4102
4·2535	1	$\frac{1}{2}$	1	×	1	$\frac{2}{3}$	$\frac{9}{10}$	= 30·0000	1·2761
1·2153	$\frac{1}{8}$	$\frac{2}{3}$	1	×	1	$\frac{2}{3}$	$\frac{9}{10}$	= 5·0000	0·1080
	$\frac{7}{8}$	$\frac{2}{3}$	1	×	1	$\frac{2}{3}$	$\frac{1}{10}$	= 3·8889	
0·0868	$\frac{7}{8}$	$\frac{2}{3}$	1	×	1	1	$\frac{1}{10}$	= 5·8333	0·0051
0·0675	$\frac{7}{8}$	$\frac{2}{3}$	1	×	1	$\frac{1}{2}$	1	= 29·1667	0·0197
0·0193	$\frac{7}{8}$	$\frac{2}{3}$	1	×	1	1	1	= 58·3333	0·0113
87·6737									47·7679

(*C*) Frequency of divergent heterocentrics among repulsion heterozygotes from matings in (*A*)

frequencies of matings giving divergent hets (%)	frequencies with which heterozygotes from (*a*) are divergent							(*a*) × (*b*)
	W-fi mate				normal mate			
	W-C′	*C′-C*	*fi-C*		*W-C′*	*C′-C*	*fi-C*	(%)	(%)
5·4688	$\frac{7}{8}$	1	1	×	1	$\frac{1}{3}$	$\frac{9}{10}$	= 26·2500	1·4356
4·2535	1	$\frac{1}{2}$	1	×	1	$\frac{1}{3}$	$\frac{1}{10}$	= 1·6667	0·0709
1·2153	$\frac{7}{8}$	$\frac{1}{3}$	1	×	1	$\frac{1}{3}$	$\frac{9}{10}$	= 8·7500	0·1080
	$\frac{1}{8}$	$\frac{1}{3}$	1	×	1	$\frac{1}{3}$	$\frac{1}{10}$	= 0·1389	
0·3906	$\frac{7}{8}$	1	1	×	1	1	$\frac{9}{10}$	= 78·7500	0·3076
0·2363	1	$\frac{1}{2}$	1	×	1	$\frac{1}{2}$	1	= 25·0000	0·0591
0·0868	$\frac{7}{8}$	$\frac{1}{3}$	1	×	1	1	$\frac{9}{10}$	= 26·2500	0·0228
0·0675	$\frac{1}{8}$	$\frac{1}{3}$	1	×	1	$\frac{1}{2}$	1	= 2·0833	0·0014
11·7188									2·0054

be given to 'H'. The F_1 CC heterozygotes were therefore mated with sibs and half-sibs, a fact which further reduces the total expectation of RD heterozygotes. Very little success in the F_3 could therefore be expected, beyond perhaps a slight deviation from 50 % favouring >50 % R in the total segregation from all F_3 matings.

(*d*) CD *heterozygotes*

A divergent coupling mating is either (i) $\frac{W\alpha}{+\beta}\frac{+\beta}{fi\alpha}$, or (ii) $\frac{W\beta}{+\alpha}\frac{+\alpha}{fi\beta}$. While a CD heterozygote requires a preparation mating in order that the centromeres remain divergent, the markers of the preparation mating must be in coupling in order that they remain in coupling in the CD heterozygote. This means that here again a short cut may be used; for preparation matings segregate for the markers and therefore some selection can be made among them for those which have the best chance of producing the CD heterozygote. Of the coupling heterozygotes being tested above for a CC one those giving independence can be mated with fidget sibs and half-sibs, and these will give a high frequency of CD heterozygotes of type (ii). Calculation by the same sort of argument as has been used in previous paragraphs gives the total expectation of CD heterozygotes in this generation as approximately 26 %. It can similarly be checked that the expectation of CC matings in the same generation by this procedure is only about 1 %.

As has been pointed out, the two repulsion heterocentrics must be obtained from preparation matings in which the markers *W* and *fi* are not both heterozygous in the same animal. This fact complicates the calculation of the frequency of these heterocentrics, and also reduces their frequencies. These obstacles can largely be overcome if a third marker is available and can be made to segregate in the preparation mating. This would be so in the present case if all repulsion preparation matings were $WfiSd\times+++$. In fact, every opportunity to make them so was taken, but they produced very few heterozygotes and none of these bred enough to be included in the final analysis.

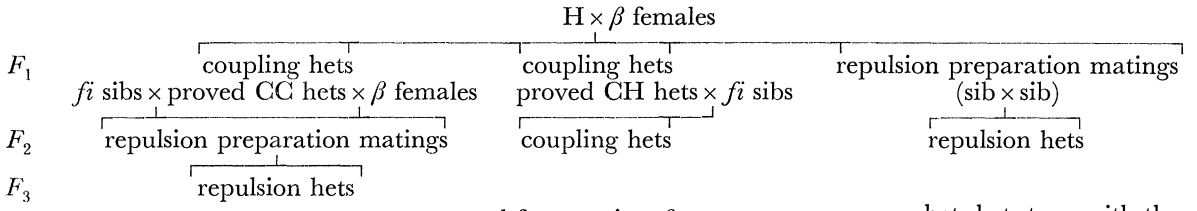
The foregoing plan for 'H''s descendants can conveniently be summarized in the form of a pedigree (figure 3).

It was decided to test at least fifteen F_1 coupling males. Further minimum numbers were decided upon on the expectation of obtaining with a 95 % certainty at least one male of each of the other three types of heterozygote. However, the difficulties already described made this impossible to achieve, and at an early stage it was decided that *all* animals of use in any part of the program should be kept and mated; a few females as well as all males possible were also tested. Even this safeguard did not produce as many data as was planned.

Although there was the general policy that each male tested should have at least five females for most of its life, any male that showed a strong deviation from 50 % recombination, in the direction expected on the basis of affinity to be the most common for its group, was given preference for extra females when there were any to spare. This was done in order that those showing quasi-linkages would do so as unequivocally as possible. Those males showing deviations in the opposite direction, that is in a direction expected to be rare or impossible, were given the same preference. This was done in the hope that if affinity were not the real explanation, there would be a reasonable amount of data on

which to build a new hypothesis. There was, therefore, no selection in any group of heterozygotes for data giving the more frequently expected deviation.

Neither was there bias in favour of heterogeneity; for preferential treatment in the assignment of mates to a particular male was only carried out after some fifty or more progeny had been bred. In order to weed out the trivial data from those worth considering, it was at first intended to include all progenies exceeding fifty; but when the experiment was completed and it was realized that the progenies were smaller than was planned, it was decided that a lower level, forty, should be taken; this decision provides an added assurance that all animals not given preferential treatment because their performance was close to independence, are included in the final analysis. One further safeguard was employed: care was taken not to terminate matings when their deviations appeared strongest; hence, as can be expected, many of the best performances did not appear as good when breeding stopped as they had done at some previous date.



group	expected frequencies of centrotypes: convergent %	expected frequencies of centrotypes: divergent %	het. het. type with the highest frequency in its group, and thus intended to be isolated there
F_1 coupling hets	19 (15)	0 (0)	(*a*) CC
F_2 repulsion hets	48 (47)	2 (2)	(*b*) RC
F_3 repulsion hets	5 or more*	40 or less*	(*c*) RD
F_2 coupling hets	1 (2)	26 (24)	(*d*) CD

* These frequencies fluctuate according to whether the proved CC hets were mated to β females, or, as an emergency measure, to *fi* sibs.
The unbracketed frequencies are those expected if H was $W\beta + \alpha/+\alpha fi\beta$.
The bracketed frequencies are those expected if H was $W\alpha + \beta/+\beta fi\alpha$.
'het' is an abbreviation for 'heterozygote', 'het het' for heterocentric heterozygote.

FIGURE 3. The planned pedigree of 'H's' descendants

IV. THE RESULTS OF THE *W*-V INVESTIGATION

(1) *The selection of* 'H'

Of several males tested for individually significant quasi-linkage, seven had bred forty-five or more progeny by the time a single outstanding performer was detected.

Table 7 gives the data from these seven animals. In general, a deviation from independence was expected to be in the direction of less than 50 % recombination, since this, as has been explained, is more readily obtained than recombination exceeding 50 % in a stock already giving quasi-linkage. However, since the centrotype of none of the parents of tested animals was known, the possibility of more than 50 % recombination was not precluded.

Accordingly the headings *N* and *R* are used, standing for those genotypes expected to be non-recombinant and recombinant respectively, on the basis of less than 50 % quasi-linkage. The pair of complementaries headed *N* are *W* and *fi*, and *R*, *W fi* and + +. Where

fi was not segregating, they are *W-Sd* and $++$, and *W* and *Sd*, respectively, except for animal 3 where these relations are reversed. The *W-fi* segregation is shown when both *Sd* and *fi* were segregating as this, rather than *W-Sd*, is expected from previous observations (table 2) to show the stronger deviation from independence. This treatment of the data is intended to ensure prominence for a good quasi-linkage, but should not bias the data in favour of heterogeneity should there be no real departures from independence. Neither does it promote the selection of an animal with a strong *W-fi* association rather than a strong *W-Sd* or *W-a*^{t} one; for in all cases the segregation of all factors was recorded, and the χ^2 value for the departures from 50 % of *W-Sd* and *W-a*^{t} not shown here, was in all cases (apart from animal 'H') less than unity.

The heterogeneity test gives a subsignificant result, p being 0·1 to 0·05. The exact probability of obtaining, out of seven trials, one animal with as high a χ^2 value as 'H''s is 1 in 16.

TABLE 7. THE SELECTION OF 'H'

tested animal	*N*	*R*	total	χ^2
(1) *W Sd* $+a^t/++$ $+a$	50	62	112	1·2857
(2) *W Sd* $+a/++$ *fi a*	49	39	88	1·1364
(3) *W* $++$ $a^t/+$*Sd* $+a$	26	19	45	1·0889
(4) *W Sd* $+a^t/++$ *fi a*	25	22	47	0·1915
(5) *W Sd* $+a/++$ *fi* a^t	26	20	46	0·7826
(6) *W Sd* $+a/++$ *fi* a^t 'H'	33	58	91	6·8681
(7) W^v $++$ $a/+$*Sd fi* a^t	46	46	92	0·0000
total χ^2	—	—	—	11·3532
totals and deviation χ^2	255	266	521	0·2322
heterogeneity χ^2				(for 6 d.f.) 11·1210
probability				*0·1 to 0·05*

(2) 'H''s *performance*

'H''s output at this stage is given in table 8*B*. His *W-fi* performance is highly significant and his *W-Sd* subsignificant.

'H' continued to breed until he was 18 months old. His total output (table 8) is also significant. χ^2 for the *W-fi* segregation is 6·2235 for 1 d.f. and for the *W-Sd*, 3·8118 for 1 d.f. It agrees with the earlier data (table 2) in that the *W-fi* segregation shows the strongest deviation from 50 %, the *W-Sd* a less strong one and the *W-a*^{t} virtually none. It is interesting to notice that the agouti locus shows no quasi-linkage in either Little's (1927) or Green's (1931) data, quoted by Michie (1953, 1955). These facts, coupled with the low probability of 1 in 16 of finding by chance as strong a quasi-linkage as 'H''s prompts the conclusion that the first part of the investigation is successful; it confirms the existence of an association of *W* with linkage group V and gives a consistent indication of the position of a point in V responsible for it.

Further, 'H''s quasi-linkage is a 'reversal' (for definition, see page 218), an observation inconsistent with chromosomal linkage, translocation and inversion, but consistent with the hypothesis of affinity. Finally, 'H''s data reinforce the otherwise small 'repulsion' data (table 2); together they show a good excess of the complementaries *W fi* and $++$ in the *W-fi* segregation, and of *W* and *Sd* in the *W-Sd*, counterbalancing the excess of the opposite pairs of complementaries in the old 'coupling' data (table 2). This balance, and the fact that the quasi-linkage ratios within the *W* class are approximately the same as

those within the non-*W* in both the *W-fi* and the *W-Sd* segregations, render very implausible all explanations in terms of unusual viability relations and of modification of the *W* penetrance.

TABLE 8. 'H' 'S PERFORMANCE

'H' = $WSd/a^t fi$ = probably $\frac{W\beta}{+\alpha}\,\frac{a+\alpha\, Sd}{a^t fi\,\beta+}$

(*A*) Total performance

	(*P*)		(*a*^{t})		(*Sd*)		(*fi*)		
	Sd	*a*^{t}*fi*	*a*^{t}*Sd*	*fi*	*a*^{t}*fi Sd*	+	*fi Sd*	*a*^{t}	totals
W	45	51	26	31	11	16	0	1	181
+	55	40	26	11	6	16	2	3	159
total	100	91	52	42	17	32	2	4	340

(*B*) Earlier and later performances

		W-a^t			*W-fi*			*W-Sd*	
	N:*R*	χ^{2*}	r.f.	*N*:*R*	χ^{2*}	r.f.	*N*:*R*	χ^{2*}	r.f.
first 91 progeny	42:49	0·5385	53·8462	33:58	6·8681	63·7363	37:54	3·1758	59·3407
later 249 progeny	125:124	0·0040	49·7992	114:135	1·7711	54·2169	115:134	1·4498	53·8153
total χ^2	—	0·5425	—	—	8·6392	—	—	4·6256	—
on total 340 progeny	167:173	0·1059	50·8824	147:193	6·2235	56·7647	152:188	3·8118	55·2941
heterogeneity χ^{2*}	—	0·4366	—	—	2·4607	—	—	0·8230	—

* The χ^2 tests here are based on an expectation of 1:1 for individual *N*:*R* and total *N*:*R*. The resulting heterogeneity χ^2 is corrected when there is a departure from 1:1, by multiplying it by $N^2/4AB$, where *A* and *B* are total *N* and total *R*, respectively, and $N=A+B$. The heterogeneity tests in succeeding tables are carried out in the same way. (This device is a simple application of the principles of the analysis of variance.)

r.f. = recombination fraction percentage.

It should perhaps be pointed out that the addition of 249 to the initial 91 mice does not alter the χ^2 values appreciably for any of the quasi-linkages. However, this discrepancy is not significant, for the heterogeneity χ^2 on the most sensitive segregation, *W-fi*, is only 2·4607 for 1 d.f. A subsignificant χ^2 may also be obtained on 'H' 's fortnightly *N*:*R* ratio for the *W-fi* quasi-linkage; it is 31·2250 for 22 d.f., $p=0{\cdot}1$ to 0·05. This suggests an age effect. Similar subsignificant effects have also been noted in other quasi-linkages, but no further discussion of the matter is intended here.

(3) *The performance of* 'H' 's *descendants*

Table 9 (appendix) gives the segregation of *W* with each of the chromosome V markers for each of the 30 tested descendants of 'H'. These fall into the four groups: F_1 coupling-convergent, F_2 repulsion-convergent, F_3 repulsion-divergent and F_2 coupling-divergent, i.e. F_1CC, F_2RC, F_3RD and F_2CD. Each group is named according to the type of quasi-linkage heterozygote which it is intended to isolate from among fellow heterozygotes most of whom are expected to show no quasi-linkage (see III, §6 and figure 3).

Tables 10, 11 and 12 analyze the data in table 9. Table 10 (appendix) gives the single-factor segregations, and tests for heterogeneity within each group of heterozygote, and table 11 (appendix) tests for single-factor heterogeneity between groups.

Table 12 gives the ratio between the two pairs of complementary genotypes for each pair of factors for each group. As explained in III, §6, a certain proportion of animals within each group is expected, on the basis of affinity, to produce a deviation from 50% recombination in one predicted direction, the direction varying for each group; the

remainder are expected to produce 50% recombination. The pair of complementaries expected to be in excess when deviations occur are labelled *G* (greater) and the opposite pair *L* (less); and this arrangement is applied to each group. As explained (§6 etc.), in groups (*a*) and (*b*) where it is hoped to isolate convergent heterocentrics, *G* and *L* coincide with *N* and *R* (non-recombinants and recombinants) respectively of chromosomal linkage, but in (*c*) and (*d*), where it is hoped to find divergent heterocentrics, these relations are reversed, the pair of complementaries expected to be greater (*G*) from the appropriate centrotype being those expected to be less (recombinant) if chromosomal linkage were operating. A heterogeneity test for the *G*:*L* ratio is given for each group.

(4) *Analysis of the performance of* 'H' 's *descendants*

Table 10 shows that within each of the four groups of heterozygote, there is a deficiency of the mutant as compared with its normal allele. Previous experience has shown that in the case of *fi* and *Sd*, this is due to impaired viability, and that in the case of *W* it is due to a small degree of imperfect penetrance (with or without viability loss). The table also shows that there is no heterogeneity for any of the single-factor ratios. Table 11, however, shows that there is heterogeneity between groups. This is very significant for the *W*:+ratio and nearly so for the *Sd*:+. The largest χ^2 value (for 1 d.f.) in the *W*:+ segregation (10·9425) is from the F_3RD group, and there are large χ^2 values throughout this group (except for a^t:*a*); the removal of these data reduces all heterogeneity χ^2 values well below significance level. That this is so is not surprising because in the *W* and *fi* segregations the frequencies of the mutants are lower here than in the other groups and that of the *Sd* mutant is also low; the third generation is the most inbred, and the single-factor ratios do appear, as can be expected, to deteriorate with inbreeding.

Thus, if the F_3 data are excluded, the remaining data are homogeneous for viability loss and imperfect penetrance. They may now be pooled and used to test each of three hypotheses. The tests are contained in table 13, which consists of three arrangements for the data in table 12. (It should perhaps be noted that, in table 13, the total χ^2 for each two-point segregation consists of the sum of the χ^2 values from each animal, and not the sum of the total χ^2 values from each group. The χ^2 values from each animal are given in table 12. The tests discern heterogeneity between individual performances and not between those of the groups.)

The source of *erratic* viability loss having been removed, the first hypothesis which suggests itself is that of *consistent* viability disturbance. Clearly a depletion of the three mutants remains after the exclusion of the F_3 data; for the *W*:+ ratio is now 1267:1306, the +:*fi* is 1073:892 and the *Sd*:+ is 994:1086. Deficiencies in both single-factor ratios in a two-point segregation sometimes disturb the equality between the complementary pairs, so mimicking the inequality produced by linkage or quasi-linkage. A test for consistent viability interaction (or indeed any other likely to disturb linkage) is provided by a heterogeneity test of this equality. Table 13*A* shows that heterogeneity in the *W*-*fi* segregation is very significant, and it is almost significant in the *W*-*Sd*. The hypothesis is therefore untenable.

Hence, there must be some real association between *W* and each of the chromosome V markers. The second rational hypothesis to be tested is, therefore, whether or not this is

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

TABLE 12. QUASI-LINKAGE SEGREGATIONS FROM 'H's' DESCENDANTS

tested animal	*W-a*^{t} G:L	*W-a*^{t} χ^2 (1:1)	*W-fi* G:L	*W-fi* χ^2 (1:1)	*W-Sd* G:L	*W-Sd* χ^2 (1:1)
			(*a*) F_1CC			
1. *a fi/a Sd W* ♂	—	—	55:65	0·8333	64:56	0·5333
2. *a fi/a*^{t} *Sd W* ♂	72:96	3·4286	86:82	0·0952	88:80	0·3810
3. *a fi/a Sd W* ♂	—	—	54:51	0·0857	54:51	0·0857
4. *a fi/a Sd W* ♂	—	—	51:31	4·8780	51:31	4·8780
5. *a fi/a Sd W* ♂	—	—	29:35	0·5625	32:32	0·0000
6. *a fi/a Sd W* ♂	—	—	72:61	0·9098	69:64	0·1880
7. *a fi/a Sd W* ♂	—	—	48:41	0·5506	81:48	8·4419
8. *a fi/a W* ♂	—	—	71:49	4·0333	—	—
9. *a fi/a W* ♂	—	—	90:112	2·3960	—	—
10. *a fi/a*^{t} *fi Sd W* ♂	155:126	2·9929	—	—	154:127	2·5943
11. *a fi/a Sd W* ♂	—	—	41:58	2·9192	40:59	3·6465
12. *a fi/a Sd W* ♀	—	—	42:29	2·3803	43:28	3·1690
13. *a fi/a*^{t} *Sd W* ♀	19:21	0·1000	19:21	0·1000	17:23	0·0900
14. *a fi/a*^{t} *Sd W* ♀	24:24	0·0000	28:20	1·3333	30:18	3·0000
15. *a fi/a Sd W* ♀	—	—	28:14	4·6667	27:15	3·4286
16. *a fi/a Sd W* ♀	—	—	17:23	0·9000	18:22	0·4000
total χ^2	—	6·5215	—	26·6439	—	31·6463
deviation χ^2	270:267	0·0168	731:692	1·0689	768:654	9·1392
(d.f.)	—	6·5047 (3)	—	25·5750 (14)	—	22·5071 (13)
corrected heterogeneity χ^2	—	6·5047	—	25·5955	—	22·6534
probability	—	*0·1 to 0·05*	—	*0·05 to 0·02*	—	*0·05 to 0·02*
			(*b*) F_2RC			
17. *a fi W/a*^{t} *Sd* ♂	52:57	0·2294	47:62	2·0642	48:61	1·5505
18. *a fi W/a*^{t} *Sd* ♂	25:25	0·0000	17:12	0·8621	29:21	1·2800
19. *a*^{t} *fi W/a Sd* ♂	69:74	0·1748	58:85	5·0979	67:76	0·5664
20. *a*^{t} *fi W/a Sd* ♂	41:45	0·1860	10:18	2·2857	39:47	0·7442
total χ^2	—	0·5902	—	10·3099	—	4·1411
deviation χ^2	187:201	0·5052	132:177	6·5534	183:205	1·2474
(d.f.)	—	0·0850 (3)	—	3·7565 (3)	—	2·8937 (3)
corrected heterogeneity χ^2	—	0·0851	—	2·8380	—	2·9030
probability	—	*>0·99*	—	*0·3 to 0·2*	—	*0·5 to 0·3*
			(*c*) F_3RD			
21. *a fi W/a Sd* ♂	—	—	24:22	0·0870	21:25	0·3478
22. *a W/a*^{t} *fi Sd* ♂	54:64	0·8475	43:33	1·3158	63:55	0·5424
23. *a fi W/a*^{t} *fi Sd* ♂	71:70	0·0071	—	—	71:70	0·0071
24. *a*^{t} *fi W/a fi Sd* ♂	85:97	0·7912	—	—	87:95	0·3516
total χ^2	—	1·6458	—	1·4028	—	1·2489
deviation χ^2	210:231	1·0000	67:55	1·1803	242:245	0·0185
(d.f.)	—	0·6458 (2)	—	0·2225 (1)	—	1·2304 (3)
corrected heterogeneity χ^2	—	0·6473	—	0·2247	—	1·2304
probability	—	*0·8 to 0·7*	—	*0·7 to 0·5*	—	*0·8 to 0·7*
			(*d*) F_2CD			
25. *a fi/a*^{t} *Sd W* ♂	32:30	0·0645	38:24	3·1613	33:29	0·2581
26. *a/a*^{t} *fi Sd W* ♂	23:27	0·3200	—	—	30:20	2·0000
27. *a fi/a fi Sd W* ♂	—	—	—	—	31:27	0·2759
28. *a*^{i} *fi/a fi Sd W* ♂	53:47	0·3600	—	—	50:50	0·0000
29. *a fi/a W* ♂	—	—	53:41	1·5319	—	—
30. *a fi/a W* ♂	—	—	43:34	1·0519	—	—
total χ_2	—	0·7445	—	5·7451	—	2·5340
deviation χ^2	108:104	0·0755	134:99	5·2575	144:126	1·2000
(d.f.)	—	0·6690 (2)	—	0·4876 (2)	—	1·3340 (3)
corrected heterogeneity χ^2	—	0·6693	—	0·4989	—	1·3400
probability	—	*0·8 to 0·7*	—	*0·8 to 0·7*	—	*0·8 to 0·7*

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

chromosomal linkage. It can readily be demonstrated that there is sufficient heterogeneity between heterozygotes to dismiss this hypothesis. Table 13*B* tests for heterogeneity in the ratios between complementary pairs arranged under the headings *N*, non-recombinant, and *R*, recombinant, pertinent to the expectations of chromosomal linkage. None of the deviation χ^2 values is significant, which suggests that the data support the hypothesis of independence rather than of chromosomal linkage. However, the very significant heterogeneity within the *W-fi* and *W-Sd* segregations renders both hypotheses implausible.

TABLE 13. TESTS FOR HETEROGENEITY IN THE RATIOS BETWEEN COMPLEMENTARY PAIRS, ARRANGED ACCORDING TO EXPECTATIONS **(*A*)**, **(*B*)**, AND **(*C*)**

	W-a^t		*W-fi*		*W-Sd*	
	Wa^t & + + : *W* & *a*^{t}	χ^2 (d.f.)	*W* & *fi* : *Wfi* & + +	χ^2 (d.f.)	*WSd* & + + : *W* & *Sd*	χ^2 (d.f.)
(*A*) viability interaction						
total χ^2		7·8562 (11)		42·6989 (22)		38·3214 (22)
deviation χ^2	572 565	0·0431 (1)	1007 958	1·2219 (1)	1099 981	6·6942 (1)
heterogeneity χ^2		7·8131 (10)		41·5019 (21)		31·7284 (21)
probability		*0·7 to 0·5*		*0·01 to 0·001*		*0·01 to 0·05*
(*B*) chromosomal linkage	*N*:*R*	χ^2 (d.f.)	*N*:*R*	χ^2 (d.f.)	*N*:*R*	χ^2 (d.f.)
total χ^2		7·8562 (11)		42·6989 (22)		38·3214 (22)
deviation χ^2	561 576	0·1979 (1)	962 1003	0·8555 (1)	1077 1003	2·6327 (1)
heterogeneity χ^2		7·6598 (10)		41·8601 (21)		35·7351 (21)
probability		*0·7 to 0·5*		*0·01 to 0·001*		*0·05 to 0·02*
(*C*) affinity	*G*:*L*	χ^2 (d.f.)	*G*:*L*	χ^2 (d.f.)	*G*:*L*	χ^2 (d.f.)
total χ^2		7·8562 (11)		42·6989 (22)		38·3214 (22)
deviation χ^2	570 567	0·0079 (1)	1024 941	3·5059 (1)	1104 976	7·8769 (1)
heterogeneity χ^2		7·8483 (10)		39·2630 (21)		30·5602 (21)
probability		*0·7 to 0·5*		*0·01 to 0·001*		*0·1 to 0·05*

The general hypothesis of affinity remains to be tested. Here, contrary to the previous cases, proof of heterogeneity supports the hypothesis, for each group is expected to be giving both quasi-linkage and independence. However, if the data are tabulated so that the complementaries are consistently arranged in the order *G*:*L*, as expected on the basis of affinity (as in table 12), heterogeneity should be slightly less than in those when chromosomal linkage is expected (table 13*B*) but the deviation χ^2 values on the total data should be greater. Table 13*C* arranges the data in this way.

All the expected differences between tables 13*B* and 13*C* are attained for the *W-fi* and *W-Sd* segregations. In particular, the very significant heterogeneity χ^2 for the *W-fi* segregation, and the very nearly significant one for the *W-Sd* (χ^2 for 21 d.f. with $p=0{\cdot}05$ is actually 32·671) are strong evidence for the hypothesis of affinity.

Closer examination of the data is nevertheless desirable before the evidence for affinity is properly evaluated. The deviations and heterogeneity within the groups of heterozygote should be, according to the way in which these heterozygotes were obtained (see III, §6), in the direction and of the magnitude expected on the basis of affinity. (The following paragraphs are a detailed examination of table 12).

(*a*) F_1 CC

The results here are very much in accordance with expectation. There is significant heterogeneity in both the *W-fi* and the *W-Sd* segregations; the deviation χ^2 for *W-Sd* is

very significant (p is 0·01 to 0·001), and the total deviations for all three two-point segregations are in the direction expected. Each of the five individual χ^2 values exceeding 3·841 ($p=0·05$) is in the expected direction. This means that there are no heterozygotes with divergent centromeres; in this group—and in this alone—a single deviation in the direction expected from divergents would have been strong evidence against affinity. Finally, the pedigree (figure 3) shows that about one in five tested heterozygotes (or one in seven, according to two alternative constitutions for their progenitor 'H') should give quasi-linkage. In fact, four (animals 4, 7, 8 and 15) out of the sixteen tested do give it, a very satisfactory agreement with expectation.

There is one unexpected point to notice. The deviation χ^2 for the *W-fi* segregation is not significant, whereas in all previous data, this pair of factors has shown stronger quasi-linkage than the *W-Sd*. The reason for this is not clear. The exclusion of animals 8, 9 and 10, which do not contribute to both segregations does not remove this anomaly although it does decrease it (the *W-fi* deviation χ^2 rises to 1·3815 and the *W-Sd* falls to 6·6337). The most probable reason is that the centromere is, in fact, almost in the middle of the *fi-Sd* segment (V, § 3); chance fluctuations in the size of the *W-fi* quasi-linkage in relation to the *W-Sd* one must therefore be expected.

(*b*) F_2RC

Here there is an apparent discrepancy. The only significant individual χ^2, that for animal 19, is for a deviation in the 'rare' direction. The pedigree (figure 3) shows that about half the tested animals should have convergent centromeres and only 2 % divergent ones. However, this is a small sample—far smaller than intended—and it is not beyond the bounds of possibility that it has included one divergent and no convergents. The lack of significant heterogeneity is probably due to the small size of the sample.

There is another reason why animal 19 could have given what can be called 'coupling data'. It is very probably a coupling convergent heterozygote. It comes from a preparation-mating intended to be *W a*^{t}*fi* male × *Sd* female. This gave 27 *W*:14 normals, an exceptionally close fit to the 2:1 expected were this mating *W* × *W* and not *W* × +; for *WW* homozygotes are lethal.* Moreover, this *W a*^{t}*fi* male, by another *Sd* female, gave 13 *W*:24 normals. This suggests that this second mating is *W* × + (when a 1:1 is expected) and that *W* is not fully penetrant. There is certainly some very small amount of misclassification in this stock, and the *Sd* mother of animal 19 was clearly a non-penetrating *W*. Animal 19 could therefore have been a coupling convergent; this is derivable with a reasonably high frequency (16·41 %) via one generation from 'H', whereas animals giving deviations from 50 % recombination in the *opposite* direction have an expected frequency less than 2 %. (It is accordingly entered in table 13*C* as a coupling convergent.)

(*c*) F_3RD

This group has been excluded from the overall analysis. Taken on its own merits, it is of very little worth, since both the *W* and the *fi* single-factor ratios are strongly disturbed (table 10). Consisting of only four segregants, it is a very small sample; however, for what

* *WW* appear anaemic and sometimes survive until 2 or 3 days old; no anaemics are recorded, but they were not looked for here and would have been dead by the time the litters were first scrutinized for classification.

it is worth, the most important two-point 'G:L' ratio (from the point of view of previous results), namely, the *W-fi* one, deviates in the expected direction. The data here do little more than fail to contradict the hypothesis of affinity.

(*d*) F_2CD

Here again, the sample is very small; so a lack of heterogeneity in all the two-point segregations is not a serious obstacle. A strong point in favour of affinity is that in all the two-point segregations the total deviations are in the expected direction, significantly so for *W-fi*; and the order of magnitude of these deviations is that expected from previous data. This is largely because animal 25 is giving quasi-linkages of the kind and magnitude expected. The data for this animal are not significant (χ^2 for *W-fi* is 3·1613), but this is not surprising since it bred only sixty-two young. It is suggested that, if affinity is operating, this animal is a double heterocentric and that, despite its subsignificant *W-fi* segregation, it should be regarded as such for mapping purposes. Being a CD heterozygote it is a repetition of 'H', a most unusual animal on any hypothesis but affinity.

(5) *Discussion and summary*

Two remaining possibilities must be discussed. The inclusion of a translocation in a program of this sort might give heterogeneous two-point segregations. It could not, however, give occasional recombination values exceeding 50 % (as do 'H' and animal 25) unless interference of the strength indicated by Fisher, Lyon & Owen (1947) and by Owen (1953*a*) for the seventh chromosome were operating, and then it could not give several values *less* than 50 % as well. It has been suggested that a translocation, and an inversion of this translocation, could do all these things. However, that this stock should contain both these anomalies (neither has been recorded in mice as occurring naturally) and that they should produce in all groups of heterozygotes just those deviations expected on the basis of affinity, seems extraordinarily improbable.

A further remark should perhaps be made. Considering that there was no way of guaranteeing the centrotype of 'H''s fidget mates, it is somewhat surprising that the data fit the expectations based on affinity as well as they do. A mate of centrotype $+\alpha fi\beta/+\beta fi\beta$, for example, could have produced divergent heterozygotes in the F_1 CC group with a high frequency, and seriously upset the expectations in the other groups.

Using the 5 % level strictly, as the basis of discrimination, it is clear that five of the twenty-six tested animals may be claimed to show quasi-linkage. Taking into account the expectation of obtaining animal 19 from an intercross for *W*, and using figure 3 for the other expectations, it may be estimated to a close approximation that the expected number is 6·36. This is a very satisfactory fit.

group	F_1C		F_2C		F_2R		
expected number	3·04	+	1·62	+	1·70	=	6·36
observed number	4	+	0	+	1	=	5

$\chi^2 = 0{\cdot}3850$ for 1 d.f. Probability $> 0{\cdot}5$.

Only two of the four types of quasi-linkage are thus demonstrated (CC and CD), failure to demonstrate the two repulsion types being explicable in view of the small numbers

available for testing. (But it should be pointed out that some of the animals whose performance is summarized in table 2 must have been RC heterozygotes.)

It should be noted that the quasi-linkage values of the five proven animals, considered together, do not allow of a linear arrangement of the four markers (table 15). This is added evidence against linkage, while the independence of *W* with the other markers is fully exemplified by the remaining twenty-one animals.

To summarize this analysis. There is very significant heterogeneity in the segregation of the complementary pairs in the two-point data from the *W*-V investigation. The tests applied and the quasi-linkages obtained contradict strongly the hypotheses of independence, of chromosomal linkage and of viability or other interactions; and they strongly support the hypothesis of affinity.

A group by group examination of the data provides further support. The segregations in the F_1CC group coincide very well with expectation. The less numerous data of the F_2CD group definitely though less strongly support the hypothesis; and this group includes, very probably, a repetition of 'H' whose data fit in with no other hypothesis considered. Data from one animal at first appear anomalous but a simple explanation may be found. The rather scanty data from the F_2RC and F_3RD groups do little more than fail to contradict the hypothesis.

Out of twenty-six animals tested altogether, the number expected to show quasi-linkage and the observed number agree satisfactorily.

In conclusion, it may be said that the data supply strong evidence for the existence of a new phenomenon, and that the hypothesis of affinity provides a very good fit.

Table 19 (appendix) summarizes the statistical evidence of the *W*-V investigation as a whole, including the earlier evidence on which it was constructed.

On the assumption that affinity is operating, the data from the investigation are now used to find the position of the centromere in chromosome V.

V. The position of the centromere in chromosome V

(1) *Criteria for the selection of data*

In contrast to chromosomal linkage experiments, the method by which the *W*-V investigation obtains affinity data does not allow of an objective decision as to which parts of the data are actually showing quasi-linkage. The course taken here is to choose a suitable significance level; ideally this would ensure that material included because the p value is lower than this level is really quasi-linkage data, and that material excluded because the p value is higher than this level is really independence data. The choice of level depends on the amount of information provided by the investigation in hand, but there is obviously no guarantee for any level that it does not result in the inclusion of a little independence data nor the exclusion of a little quasi-linkage data.

The *W*-V investigation clearly involves rather loose quasi-linkages and consists of progenies of very variable size. It is suggested, therefore, that the 5 % level is too low and might exclude small but informative bodies of data; the 10 % level is preferred ($\chi_2 = 2{\cdot}706$). (Table 12 is used for this selection.)

A second problem arises: are data for *all* two-point quasi-linkages with χ^2 exceeding 2·706 to be included? In general, it is clear that data with χ^2 exceeding 2·706 from *W-fi* or *W-Sd* or both must be accepted, otherwise there will be bias in the location of the centromere in relation to *fi* and *Sd*. Moreover, animals which segregate for *either fi or Sd* (but not both) must for this purpose be excluded. But as regards the *W-a*^{t} segregations, a further selection of the data must be made, and a comparison with previous data used as the criterion; otherwise a heterozygote such as animal 2, whose *W-a*^{t} χ^2 value slightly exceeds 2·706 and is in the direction expected with frequency zero, and whose *W-Sd* and *W-fi* deviations are much less strong, would have to be included. Since *W-a*^{t} did not show strong quasi-linkage in the earlier data (table 2 and 'H', table 8), χ^2 values exceeding 2·706 for this pair of factors must be ignored. (This decision does in fact affect only animal 2.)

One other body of data must, on this kind of argument, be excluded; that from animal 11, both of whose *W-Sd* and *W-fi* deviations from independence are in the direction expected with zero frequency. It should be noted that the χ^2 values here do not exceed 3·841 and that if as low a level as 2·706 is used, the need for discrimination against such data—whose deviations are probably due to chance—must be expected.

Conversely, no animal whose χ^2 value for *W-fi* or *WSd* greatly exceeds 2·706 should be excluded, even if there is some doubt as to its real constitution. Animal 19, by the standards set above, is certainly giving quasi-linkage; whether it is an RD or CC heterozygote (see IV, §4) makes no difference to the way in which its data are used for mapping purposes. This animal is therefore included.

Finally, quasi-linkage data are not included from heterozygotes whose chromosomal linkage data are anomalous. A test for heterogeneity within the linkage data for the eight animals conforming to the standards set above for quasi-linkage segregations calls for the exclusion of animal 12; for the test shows that there is significant heterogeneity only in the *fi- Sd* segregation and that animal 12, whose recombination fraction of 7·04 is very much below the average, is mainly responsible.

The performance of the seven animals so selected is not only homogeneous as regards chromosomal linkage, but, although not balanced, is also in agreement with the large body of data from the balanced three-point backcross experiment for *a*^{t}, *fi* and *Sd*; for the latter data, from male and female heterozygotes combined (Wallace 1957) gives the *a*^{t}*-fi* value as 31·48 ± 1·40 %, the *fi-Sd* as 22·29 ± 1·26% and the *a*^{t}*-Sd* as 44·13 ± 1·50 %. The seven *W*-V animals give these values as 29·51 ± 1·87 %, 19·85 ± 1·41 % and 41·82 ± 2·03 % respectively. (This is, incidentally, good evidence that there is nothing unusual about their meiotic behaviour, and that the cause of their quasi-linkages is not itself a disruptive agent; it is, in fact, indirect evidence that affinity, which is not expected to disturb meiosis except in the one definitive interchromosomal way, is the agent.) Despite this agreement, the chromosomal linkage data from the three-point backcross are used for mapping purposes, in preference to that from the affinity animals, as being the most accurate available.

(2) *The data selected*

Table 14 (appendix) gives the simultaneous segregations of the factors *W*, *a*^{t}, *fi* and *Sd* from the seven 'affinity' animals.

It may be verified from this that there is significant inequality in the $+:fi$ ratio and a nearly significant one in the $W:+$. (The $Sd:+$ one is very insignificant.) Some sub-significant viability disturbance may thus be expected in the *W-fi* quasi-linkage. It is, therefore, worth noticing that there *is* good balance here, for the coupling data total 402 and the repulsion 404; so that estimates of quasi-linkage values derived by addition from the whole data may be expected to have virtually no bias.

It is advisable, however, to consider the output of the two females separately from that of the males, for there is an almost significant sex difference in the *fi-Sd* recombination value as shown by the three-point backcross, and it may be that separation values also vary between the two sexes. This subtraction of the female from the male data does not appreciably mar the balance, because the former constitute only about one-ninth of the whole.

Table 15 summarizes these data in a concise form for mapping purposes.

TABLE 15. A SUMMARY OF THE SIMULTANEOUS SEGREGATIONS FROM THE SEVEN 'AFFINITY' HETEROZYGOTES

(I) Male gametogenesis

total data ignoring a^t locus	(*P*) 341		(*W*) 236		(*S*) 74		(*fi*) 65		totals 716
total data including a^t locus	(*P*) 171	(a^t) 84	(*W*) 132	(Wa^t) 55	(*Sd*) 48	(a^tSd) 13	(*fi*) 5	(a^tfi) 37	545

(II) Female gametogenesis

total data ignoring a^t locus	(*P*) 46		(*W*) 23		(*Sd*) 10		(*fi*) 11		90
total data including a^t locus	(*P*) 13	(a^t) 10	(*W*) 8	(Wa^t) 5	(*Sd*) 4	(a^tSd) 1	(*fi*) 2	(a^tfi) 5	48
				W-a^t		*W-fi*		*W-Sd*	
recombination values on combined data				48·9039 %		41·5633 %		42·5558 %	

N.B. All values exceeding 50 % $(1-y)$ are expressed as less than 50 % (i.e. as y).

(3) *The centromere map (given in recombination values)*

(*a*) *The map for male gametogenesis*

Using the following values, derived directly from table 15:

$$W\text{-}fi = 301/716, \quad W\text{-}Sd = 310/716 \quad \text{and} \quad fi\text{-}Sd = 24{\cdot}0437\,\%,$$

the latter being the observed value for males in the three-point linkage backcross; and substituting them in equations (6) and (7) (see III, §3), the estimates of *fi-C* and *Sd-C* become

$$fi\text{-}C = 10{\cdot}7424\,\%, \quad Sd\text{-}C = 16{\cdot}9410\,\%.$$

This suggests the order *fi-C-Sd*. Confirmation that this is correct, rather than the order *C-fi-Sd* is obtained by using the a^t locus. With

$$W\text{-}a^t = 266/545 \quad \text{and} \quad a^t\text{-}Sd = 43{\cdot}1694\,\%$$

(observations derived respectively from table 15 and the three-point backcross), and the additional equation

$$(1-2a^t\text{-}C)^2 = (1-2W\text{-}a^t)\,(1-2a^t\text{-}Sd)/(1-2W\text{-}Sd) \qquad (11)$$

derived by the same argument that provides equations (6) and (7), the estimate for a^t-C becomes

$$a^t\text{-}C = 42{\cdot}2051\ \%.$$

As expected, it is slightly less than the a^t-Sd value, and more than the a^t-fi. This is a good confirmation of the order fi-C-Sd; for if the order were a^t-C-fi-Sd, the a^t-C value would be smaller than the a^t-fi.

It should perhaps be stressed that these data do not provide much information on the W-a^t quasi-linkage value, and that this may be biased unavoidably by the method used to isolate the data showing 'affinity'. Ignoring this locus, there are only three observable variables: $(1-2W\text{-}Sd)$, $(1-2W\text{-}fi)$ and $(1-2fi\text{-}Sd)$, from which the three parameters $(1-2fi\text{-}C)$, $(1-2Sd\text{-}C)$ and $(1-2W\text{-}C')$ $(1-2C\text{-}C')$ may be estimated. Thus there are no degrees of freedom for measuring goodness of fit, and hence the location of the centromere outside the segment fi-Sd and on the fi side cannot be completely excluded by these 'affinity' data alone.

(*b*) *The map for female gametogenesis*

Using the following values

$$W\text{-}fi = 34/90, \quad W\text{-}Sd = 33/90 \quad \text{and} \quad fi\text{-}Sd = 20{\cdot}5454\ \%$$

derived as for males, and the same equations, the estimates of fi-C and Sd-C become

$$fi\text{-}C = 13{\cdot}2575\ \%, \quad Sd\text{-}C = 9{\cdot}9174\ \%.$$

The a^t-C value, if calculated as for males, would be infinity; this is because the observed W-a^t value is exactly 50 %. If this value were based on a large observation, it would favour the order a^t-fi-Sd-C; but since it is based on only forty-eight progeny, it does no more than contradict the order a^t-C-fi-Sd.

(*c*) *The combined map*

Since this is sometimes useful for comparison with other chromosomes, it is given here. The following values are used

$$W\text{-}fi = 335/806, \quad W\text{-}Sd = 343/806 \quad \text{and} \quad fi\text{-}Sd = 23{\cdot}6550\ \%.$$

The latter is derived from the three-point back-cross value for females and for males in the ratio 1 for female to 8 for male, according to the ratio 1:8 of the total data for females (90 progeny) to that for males (716) in table 15. Using equations (6) and (7) as before, the estimates for fi-C and Sd-C become

$$fi\text{-}C = 11{\cdot}3622\ \%, \quad Sd\text{-}C = 15{\cdot}9078\ \%.$$

Using the observed value 290/593 for W-a^t and the value 43·3829 for a^t-Sd, the latter derived in the same way as the fi-Sd, the estimate of a^t-C is obtained

$$a^t\text{-}C = 43{\cdot}0202\ \%.$$

(4) *The third chromosome (III)*

If the relevant values given in §3 (*a*) and (*b*) are used in equation (8), the following estimates of the compound $(1-2W\text{-}C')^2\ (1-2C'\text{-}C)^2$ are obtained: for males 0·0411, and for females 0·1107. These give the maximum values for *either* W-C or C'-C; for

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

if W-C' is assumed to be 0, then these values correspond to $(1-2C'\text{-}C)^2$, and conversely, if C'-C is assumed to be 0, these values correspond to $(1-2W\text{-}C')^2$. Thus, for males the maximum value of W-C' *or* C'-$C = 39{\cdot}86\,\%$; for females, $33{\cdot}37\,\%$.

Thus, the frequency of recombination of W with its centromere is probably not more than about 40 %, and the average centromere separation value in this experiment is also not more than about 40 %.

(5) *Evidence concerning the centromere from interference relations*

If it is conceded that 'affinity' is operating in the W-V investigation, then the points located in the maps above are in fact the centromeres. If, however, the operative points are not the centromeres, then the evidence obtained in this investigation still indicates that they have the qualities here attributed to the centromeres, namely, that they can be of at least two different types having something in common with homologous points on other chromosomes such that similar ones attract, and that these differences are permanent. If these points are not the centromeres, then some term other than 'affinity' is required to describe the phenomenon.

There is some relevant evidence from another source. Wallace's balanced three-point backcross using the markers a^t, *fi* and *Sd* provides the following K values: for males $0{\cdot}7113 \pm 0{\cdot}1729$, for females, $0{\cdot}8493 \pm 0{\cdot}1584$ and for the combined data $0{\cdot}7863 \pm 0{\cdot}1168$. The male and female values are not significantly different from each other, neither are the separate and combined values significantly different from unity.

K is a measurement of interference based on the work of Kosambi (1944), which was introduced by Sir Ronald Fisher and developed by Owen (1950, 1951, 1953*b*). Assuming that there is no interference across the centromere, K is expected to rise as the centromere is approached. On the particular interference metric developed by Owen, the $\frac{1}{4}\chi^2_4$, which he has found (1948) to fit well all the available *Drosophila* data, K is expected to be unity for segments between the centromere and terminus, to decline near the terminus and to exceed unity for segments enclosing the centromere. The K values for a^t, *fi* and *Sd*, therefore, suggest that a^t-*Sd* does not enclose the centromere.

However, these values are not entirely inconsistent with the location of the centromere between *fi* and *Sd*; for there are three bodies of linkage evidence which indicate that it is at the *Sd* end of the chromosome, and one argument against it being beyond *Sd*.

From a balanced backcross experiment, Owen (1953*b*) obtained for the segment *a-un-we-pa* the following K values: for males $0{\cdot}9663 \pm 0{\cdot}6829$, for females $0{\cdot}1624 \pm 0{\cdot}1641$, and combined $0{\cdot}2092 \pm 0{\cdot}1509$. The combined value is thus lower than that for a^t-*Sd* and is clearly significantly less than unity. From a further balanced backcross for the segment *Ra*-a^t-*we*, Parsons (1958) has obtained for males a K value of $0{\cdot}354 \pm 0{\cdot}157$ and for females $0{\cdot}456 \pm 0{\cdot}172$.*

These two bodies of data thus suggest that the segment *Ra-pa* is near the terminus and that the centromere is therefore at the *Sd* end of the arm.

* The linkage group V map, in recombination percentages, for the factors mentioned here is:

for females	20·6	4·7	7·4	4·2	17·0	20·5	
	Ra——	*a*^{t}——	*un*——	*we*——	*pa*——	*fi*——	*Sd*
for males	23·3	4·6	4·4	2·2	23·0	24·0	

Borger (1950) determined the relation of the series *a*^{t}-*fi*-*Sd* to *a*^{t}-*pa*, establishing the order *a*^{t}-*pa*-*Sd*. His three-point backcross using these factors which was not designed to give a quantitative map, was not balanced and the progenies reared were not large (total 277). His information on interference is therefore less reliable than Wallace's, Owen's and Parsons', but the *K* value derivable from it is suggestive. At $0{\cdot}9456 \pm 0{\cdot}316$, it is not significantly different from unity, and is thus consistent with a mid-arm location of these segments and thus with the centromere being at the *Sd* end.

The argument that it is not beyond *Sd* is as follows. The distance between *a*^{t} and *Sd* is about 60 cM, as calculated from the *a*^{t}-*fi*-*Sd* backcross, assuming $K=1$. Carter and Phillips (1954) have shown that *Ra* recombines with *a*^{t} with a frequency $24{\cdot}2 \pm 2{\cdot}5$ %, and that it is beyond *a*^{t}, and Parsons has confirmed this. If it is conceded that *Ra* is near the terminus, and if Owen's $\frac{1}{4}\chi^2_4$ metric is assumed, then at least 30 cM must be allowed for the *Ra*-*a*^{t} interval, and some further unspecifiable distance between *Ra* and the terminus. Thus, if terminus-*Sd* is all one arm, it must greatly exceed 90 cM. If the centromere in chromosome V is median, and the chromosome is of average length, then only 60 cM are expected for each arm. (This is the average on the basis of chiasmata counts, Crew & Koller 1932; Slizynski 1949): thus an arm greatly exceeding 90 cM is intolerable on this hypothesis. It seems, therefore, that either the centromere here is nearly terminal, in which case it may be beyond *Sd*, or it is median, in which case its location between *fi* and *Sd* is consistent.

Thus, it may be said that an interpretation of the results of linkage experiments based on Owen's development of the Kosambi relations, is not opposed to the location of the centromere between *fi* and *Sd*.

(6) *Cytological evidence*

No doubt it will eventually become possible to verify cytologically the position of the centromere in this and other chromosomes. It is, however, difficult to know what reliance may be placed on the present somewhat conflicting evidence. Slizynski's map (1949) contains at least six of the twenty chromosome pairs with about median centromeres whereas Sachs (1955) states that 'the mitotic metaphase, when the centromeres can be clearly seen, shows no chromosome with a median or submedian centromere'.

(7) *A tentative interference map of the centromere*

The exact positions of the supposed centromere for males and females given by the *W*-V investigation may now be used, in conjunction with the observed recombination values obtained from linkage experiments, to obtain a map in metrical units based on Owen's model of a $\frac{1}{4}\chi^2_4$ interference metric.

(*a*) *Male map*

It can be verified (by a method outlined in the Introduction to Fisher & Yates (1953)) that the trial values

$$T\text{—}34\text{—}a^t\text{—}33\text{—}fi\text{—}28\text{—}C \quad \text{(in metrical units)},$$

(where *T* is the terminus and *C* the centromere) yield the recombination values *fi*-*C*, 10·8871 % and *a*^{t}-*fi*, 27·2801 %; these are a reasonable fit to the observed values 10·7424 % (from the *W*-V investigation) and 27·1403 % (from Wallace's three-point backcross).

Using the observed *fi-Sd* value 24·0437% (also from Wallace's three-point backcross), a value for *Sd-C* of 16·8187% is obtained, again sufficiently close to the value 16·9410% obtained by the *W*-V investigation. The trial metrical values above, then, are a good fit to the observed recombination values.

To test whether they reflect the interference relations for these segments, it is necessary to find the frequencies of the modes of gamete formation for the segments *a*^{t}*-fi-Sd*, and to compare them with the observed frequencies (of Wallace's three-point backcross). See table 16.

TABLE 16. TEST FOR THE GOODNESS OF FIT OF THE TRIAL VALUES OF A METRICAL MAP WITH OBSERVED RECOMBINATION VALUES—MALE GAMETOGENESIS

modes	expected	observed	$(a\text{-}mn)^2/mn$
(*P*)	300·5851	290	0·3728
(*a*^{t})	116·4149	127	0·9624
(*Sd*)	98·6470	110	1·3066
(*fi*)	33·3528	22	3·8643
	549·0000	549	$\chi^2 = 6{\cdot}5061$ for 1 d.f.
			probability *0·02 to 0·01*

This is a poor fit, though not absolutely remote. Clearly the greatest part of the discrepancy is concentrated in the paucity of observed double recombinants compared with those expected. This is reflected in a comparison of the *K* values derived from these expectations and from the observed frequencies, namely $K = 1{\cdot}1822$ and $K = 0{\cdot}7113 \pm 1729$, respectively. It is apparent, therefore, that although better trial values for the metrical values might improve the fit, the two *K* values still would not coincide.

(*b*) *Female map*

The trial values

$$T\text{—}22\text{—}a^t\text{—}42\text{—}fi\text{—}31\text{—}C$$

yield the recombination values *fi-C*, 12·8281% and *a*^{t}*-fi*, 36·1876%; these are a reasonable fit to the observed values 13·2575% and 35·8182% (from the *W*-V investigation and Wallace's three-point backcross, respectively). Using the observed *fi-Sd* value 20·5454% (also from the latter source), a value for *Sd-C* of 10·3806% is obtained, again sufficiently close to the value 9·9174% obtained by the *W*-V investigation.

A test of the goodness of fit of these trial values and the total observations for the segments *a*^{t}*-fi-Sd*, is shown in table 17

TABLE 17. TEST FOR THE GOODNESS OF FIT OF THE TRIAL VALUES OF A METRICAL MAP WITH OBSERVED RECOMBINATION VALUES—FEMALE GAMETOGENESIS

modes	expected	observed	$(a\text{-}mn)^2/mn$
(*P*)	274·7503	271	0·0512
(*a*^{t})	162·2497	166	0·0867
(*Sd*)	76·2174	82	0·4387
(*fi*)	36·7826	31	0·9091
	550·0000	550	$\chi^2 = 1{\cdot}4857$ for 1 d.f.
			probability *0·3 to 0·2*

The fit is very good. Here again, the discrepancy, small as it is, is mainly in the paucity of observed double recombinants compared with those expected. Thus, the *K* value derived

from these expectations, $K=1{\cdot}0373$ exceeds that derived from the observed frequencies, $K=0{\cdot}8493\pm0{\cdot}1584$ as in the case of male gametogenesis, but here the two K values do not differ significantly. It seems, therefore, that the trial metrical values above give an extremely good fit on the available evidence.

It should be pointed out that this evidence is not as abundant as for males; the progenies reared from each sex in the three-point linkage backcross are equal, but only ninety progeny were reared from females contributing to mapping evidence in the *W*-V investigation, whereas 716 were reared from males.

(*c*) *Alternative male map*

The good fit obtained for females suggests that the location of the centromeres nearer *Sd* than *fi* may improve the fit for males.

The trial metrical values

$$T\text{—}22\text{—}a^t\text{—}29\text{—}fi\text{—}34\text{—}C$$

produce the goodness of fit shown in the following table (table 18):

TABLE 18. TEST FOR THE GOODNESS OF FIT OF THE TRIAL VALUES OF AN ALTERNATIVE MAP WITH OBSERVED RECOMBINATION VALUES—MALE GAMETOGENESIS

mode	expected	observed	$(a\text{-}mn)^2/mn$
(*P*)	298·5777	290	0·2464
(a^t)	118·4223	127	0·6213
(*Sd*)	100·6842	110	0·8619
(*fi*)	31·3158	22	2·7713
	549·0000	549	$\chi^2=4{\cdot}5009$ for 1 d.f.
			probability *0·05 to 0·02*

These trial values do, as expected, give a much improved fit, though not yet perfect. The recombination values derived from them are

$$T\text{—}31{\cdot}7430\,\%\text{—}a^t\text{—}27{\cdot}2747\,\%\text{—}fi\text{—}15{\cdot}0519\,\%\text{—}C\text{—}12{\cdot}8645\,\%\text{—}Sd.$$

It should be noted that part of the improvement is due to a decrease in the assumed metrical value between a^t and the terminus. The recombination value it provides, 31·743 %, is consistent with the *Ra*-a^t value of $23{\cdot}35\pm1{\cdot}48$ % from Parsons's three-point experiment. His data from females provide the value $20{\cdot}56\pm1{\cdot}41$ %, again well below the 32·658 % allowed for that sex (§7) in the interval from a^t to the terminus.

That the quasi-linkage values obtained from the *W*-V investigation, though favouring the location of the centromere nearer to *fi* than to *Sd*, are not inconsistent with this shift in the opposite direction, can be seen from the appropriate χ^2 test. This is obtained by the following argument. It follows from equations (6) and (7) (III, §3), that

$$\frac{1-2Sd\text{-}C}{1-2fi\text{-}C}=\frac{1-2W\text{-}Sd}{1-2W\text{-}fi}. \qquad (12)$$

This relation provides a criterion by which any chosen values of *Sd*-*C* and *fi*-*C* may be compared with those expected from affinity data.

If values are assigned to the unobservable recombination fractions *Sd-C* and *fi-C* in accordance with the map to be tested, the ratio

$$\frac{1-2Sd\text{-}C}{1-2fi\text{-}C}=\lambda \tag{13}$$

may be calculated. This has a known value, when used as in equation (12), and gives a linear relation between the frequencies of the four modes of gamete formation observable in the affinity data. Thus if a, b, c, d correspond with the observed frequencies

(P)	(W)	(Sd)	(fi)	
341	236	74	65	respectively (given in table 15),

then

$$\frac{1-2W\text{-}Sd}{1-2W\text{-}fi}=\frac{a-b-c+d}{a-b+c-d}=\frac{96}{114}$$

and this ratio will be the value assigned to λ.

The relation

$$\lambda=\frac{a-b-c+d}{a-b+c-d}$$

may equally be written

$$(\lambda-1)\,a-(\lambda-1)\,b+(\lambda+1)\,c-(\lambda+1)\,d=0, \tag{14}$$

implying in this form that a linear function of the frequencies observable should be zero for each value of λ in terms of which the coefficients of the function are expressed. For any arbitrarily chosen value of λ the observed frequencies will show some discrepancy, and the significance of this may be easily determined, since the sampling variance of any linear function of observable, mutually exclusive, frequencies is always known (Fisher 1954, p. 307).

In general, the sampling variance of any such linear function of observable frequencies, the expected value of which is zero, is expressible in terms of the expectation, so that

$$\mathrm{V}\,(14)=(\lambda-1)^2\,(\alpha+\beta)+(\lambda+1)^2\,(\gamma+\delta),$$

where α, β, γ, δ stand for the values of a, b, c, d expected.

Hence

$$\frac{\{(a-b)\,(\lambda-1)+(c-d)\,(\lambda+1)\}^2}{(\lambda-1)^2\,(\alpha+\beta)+(\lambda+1)^2\,(\gamma+\delta)} \tag{15}$$

is distributed as χ^2 for 1 degree of freedom. In this expression the marginal values $(a+b)$ and $(c+d)$ may be substituted for the expectations $(\alpha+\beta)$ and $(\gamma+\delta)$. In the present case, where *Sd-C* and *fi-C* are assumed to be 12·865 % and 15·052 %, respectively, χ^2 becomes 1·067, clearly insignificant.

The test (15) may be used, with values of χ^2 corresponding to chosen significance levels, to find the limits of λ tolerated by the affinity data at those levels. Using relation (13) and

$$(1-2fi\text{-}Sd)=(1-2fi\text{-}C)\,(1-2Sd\text{-}C) \tag{5}$$

the limits of the values *Sd-C* and *fi-C* may then be determined.

Thus, when a χ^2 value of 3·841 (for the 5 % level) is used in the test (15), the limiting values of λ are 0·47 and 1·34.

The value of λ when the centromere is assumed to be exactly at the *fi* locus, is found by using equation (13); zero is substituted for the recombination fraction *fi*-*C*, and the value for *fi*-*Sd* as found from table 15 is substituted for the recombination fraction *Sd*-*C*. This is 0·62. Since the lower limiting value 0·47 is clearly less than this, it may be concluded that the affinity data tolerate, at the 5 % level, a position of the centromere beyond *fi*. However, they do not tolerate a position beyond *Sd*, for the limiting position for the 5 % level (when $\lambda = 1{\cdot}34$) is as follows:

fi—16·16 %—*C*—4·81 %—*Sd*.

A map equally consistent with the three-point backcross data and that from the *W*-V investigation can be obtained on the present system of mapping, by the minimization of the summed χ^2 value appropriate to the fit tested above and that tested in table 18. The *W*-V data, the first of its kind, is not extensive enough to warrant further calculation and adjustment beyond the remark that the centromere position most consistent with both bodies of data is probably somewhat nearer *Sd* than is indicated by the alternative male map.

(8) *Conclusion*

It is clear (from § 3) that the *W*-V investigation provides strong evidence that the position of the point in chromosome V responsible for its association with chromosome III is located between *fi* and *Sd*.

Qualitative agreement with the hypothesis that the centromere is at the *Sd* end is obtained (in § 5) from consideration of the observed interference value, *K*, calculated from Wallace's three-point backcross involving a^t, *fi* and *Sd*, and from three other multi-point backcrosses whose *K* values (though in one case less precise) are worth considering. An argument, based on the known average lengths of the chromosomes, is given which shows that unless the centromere is subterminal, its location beyond *Sd* is unlikely. Cytological evidence is conflicting upon the question of whether any or all of the centromeres are subterminal.

An attempt is made to obtain *quantitative* agreement with the hypothesis that the centromere is in fact the point between *fi* and *Sd* located by the *W*-V investigation. Metrical maps based on recombination values from both the *W*-V investigation and Wallace's linkage backcross, are constructed on the assumption of Owen's $\frac{1}{4}\chi^2_4$ interference metric. Here the agreement is not conclusive; for the fit of a metrical map based on the exact location of the centromere indicated by the *W*-V data from the males, is significantly poor though not absolutely remote; while the fit of such a metrical map to the female body of data, which is considerably smaller than that of the males, is very good. An alternative map for males, assuming a slight shift in the position of the centromere so that it coincides more closely with that indicated by the female data, provides a better fit. An argument is put forward showing that further adjustment would produce a map equally consistent with the three-point data and the *W*-V data, whose fit to both would not be significantly poor.

This result is somewhat surprising when it is considered that two rather stringent assumptions have been made—first, that there is no interference across the centromere, whereas there is no decisive evidence that there is not a small amount; and second, that the $\frac{1}{4}\chi^2_4$ metric, which indeed provides a reasonable reflexion of the operation of interference in *Drosophila* is a perfect one for the mouse.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Further data both upon interference and upon quasi-linkage values would provide more conclusive evidence on all these points. A five-point program, involving *Ra*, a^t, *we*, *fi* and *Sd*, is now under construction, and material for further affinity work is also being developed. The former will supply several *K* values which should indicate the position of the centromere more clearly than have previous experiments. The latter will give more accurate data than the present *W*-V investigation as to the position of the point in V responsible for the association with it of the chromosome III markers.

For the present it seems reasonable to assume, until it is proved otherwise, that the centromere is the causative agent for quasi-linkage.

The following maps are proposed as fitting most closely all the available evidence (figure 4). The recombination values are derived from the metrical values as shown in § 6, and the map distances from Owen's (1950) formula, which may be written

$$x = \tfrac{1}{4}(\cosh 4u - 1) - \tfrac{1}{4}\tanh 2T(\sinh 4u - 4u), \qquad (16)$$

where *u* is the metrical distance from marker to centromere and *T* the metrical length of the arm. No direct estimation can be made of the length of the *Sd-C* segment, but an upper limit is set by the assumption of no interference, that is by Haldane's (1919) relation $x = \frac{1}{2}\ln(1-2y)$, where *x* is the map length and *y* the recombination value. A lower limit is set by the recombination value itself.

♀♀	*T*	a^t	*fi*	*C*	*Sd*	arm length	
	32·658	36·188	12·828	10·381			recombination values
	34·576	42·057	14·210	11·636		90·843	map distance
	22	42	31			95	metrical distance
♂♂	*T*	a^t	*fi*	*C*	*Sd*	arm length	
	31·743	27·275	15·052	12·865			recombination values
	33·598	29·753	16·159	14·872		79·510	map distance
	22	29	34			85	metrical distance

FIGURE 4. The map of chromosome V

I wish to thank Sir Ronald Fisher, F.R.S. for his constant encouragement, and for his direction in the subject of mapping. I am also indebted to Dr Jane Brandt and Mr P. A. Parsons for their help in checking calculations, and to Mr S. A. Mallyon for very competent technical assistance.

REFERENCES

Borger, R. 1950 Order of genes in the fifth linkage group of the house mouse. *Nature, Lond.* **166**, 697.

Carter, T. C. & Phillips, R. J. S. 1954 Ragged, a semi-dominant coat texture mutant in the house mouse. *J. Hered.* **45**, 150–154.

Crew, F. A. E. & Koller, P. C. 1932 The sex incidence of chiasma frequency and genetical crossing-over in the mouse. *J. Genet.* **26**, 359–383.

Fisher, R. A. 1954 *Statistical methods for research workers.* 12th ed. London: Oliver and Boyd.

Fisher, R. A., Lyon, M. F. & Owen, A. R. G. 1947 The sex chromosome in the house mouse. *Heredity*, **1**, 355–365.

Fisher, R. A. & Yates, F. 1953 *Statistical tables for biological, agricultural and medical research.* 4th ed. London: Oliver and Boyd.

Gates, W. H. 1926 The Japanese waltzing mouse: its origin, heredity and relation to the genetic characters of other varieties of mice. *Publ. Carneg. Instn*, no. 337, 83–138.

Green, C. V. 1931 Size inheritance and growth in a mouse species cross (*Mus musculus* × *Mus bactrianus*). II. Birthweights. *J. Exp. Zool.* **58**, 247–258.

Grüneberg, H. 1943–52 *The genetics of the mouse.* London: Cambridge University Press, and *Bibliographia Genetica* XV.

Haldane, J. B. S. 1919 The combination of linkage values and the calculation of distances between the loci of linked factors. *J. Genet.* **8**, 299–309.

Kosambi, D. D. 1944 The estimation of map distances from recombination values. *Ann. Eugen., Lond.* **12**, 172–175.

Little, C. C. 1927 Notes on a species cross in mice and on a hypothesis concerning the quantitative potentiality of genes. *Science*, **66**, 542–543.

Michie, D. 1953 Affinity: a new genetic phenomenon in the house mouse. *Nature, Lond.* **171**, 26–27.

Michie, D. 1955 'Affinity.' *Proc. Roy. Soc.* B, **144**, 241–259.

Owen, A. R. G. 1948 *The theory of genetical recombination* (etc.). Ph.D. Dissertation. Cambridge: University Library.

Owen, A. R. G. 1949 The theory of genetical recombination. I. Long chromosome arms. *Proc. Roy. Soc.* B, **136**, 67–94.

Owen, A. R. G. 1950 The theory of genetical recombination. *Advanc. Genetics* **3**, 117–157.

Owen, A. R. G. 1951 An extension of Kosambi's formula. *Nature, Lond.* **168**, 208.

Owen, A. R. G. 1953*a* Super-recombination in the sex chromosome of the mouse. *Heredity*, **7**, 103–110.

Owen, A. R. G. 1953*b* The analysis of multiple linkage data. *Heredity*, **7**, 247–264.

Parsons, P. A. 1958 Additional three-point data for linkage group V of the house mouse. *Heredity* (in the Press).

Sachs, L. 1955 The possibilities of crossing-over between the sex chromosomes of the mouse. *Genetica*, **27**, 309–322.

Schwarz, E. & Schwarz, H. K. 1943 The wild and commensal stocks of the house mouse, *Mus musculus* Linnaeus. *J. Mammal.* **24**, 59–72.

Slizynski, B. M. 1949 A preliminary pachytene chromosome map of the house mouse. *J. Hered.* **49**, 242–245.

Trow, A. H. 1913 Forms of reduplication: primary and secondary. *J. Genet.* **2**, 313–324.

Wallace, M. E. 1950 Locus of the gene 'fidget' in the house mouse. *Nature, Lond.* **166**, 407.

Wallace, M. E. 1953 Affinity: a new genetic phenomenon in the house mouse. *Nature, Lond.* **171**, 27–28.

Wallace, M. E. 1954*a* Affinity: a new genetic phenomenon observed in the house mouse. *Proc. IXth Int. Congr. Genetics* (*Bellagio* 1953), published as supplement to *Caryologia*, **6**, 1048–51.

Wallace, M. E. 1954*b* *Studies in mouse genetics.* Ph.D. Dissertation. Cambridge: University Library.

Wallace, M. E. 1957 A balanced three-point experiment for linkage group V of the house mouse. *Heredity*, **11**, 223–258.

Wright, M. E. (=Wallace) 1947 Two sex-linkages in the house mouse with unusual recombination values. *Heredity*, **1**, 349–354.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Appendix

Table 9. Segregation of *W* with V markers in 'H''s descendants

tested animal	*W* *a*^{t}	*W* *a*	+ *a*^{t}	+ *a*	total	*W* +	*W* *fi*	+ +	+ *fi*	total	*W* *Sd*	*W* +	+ *Sd*	+ +	total
						(*a*) F_1CC									
1. *a fi/a Sd W* ♂	—	—	—	—	—	30	26	39	25	120	32	24	32	32	120
2. *a fi/a*^{t} *Sd W* ♂	31	51	45	41	168	41	41	41	45	168	40	42	38	48	168
3. *a fi/a Sd W* ♂	—	—	—	—	—	24	26	25	30	105	23	27	24	31	105
4. *a fi/a Sd W* ♂	—	—	—	—	—	26	9	22	25	82	27	8	23	24	82
5. *a fi/a Sd W* ♂	—	—	—	—	—	13	12	23	16	64	11	14	18	21	64
6. *a fi/a Sd W*♂	—	—	—	—	—	39	30	31	33	133	35	34	30	34	133
7.* *a fi/a Sd W* ♂	—	—	—	—	—	23	16	25	25	89	40	24	24	41	129
8. *a fi/a W* ♂	—	—	—	—	—	38	20	29	33	120	—	—	—	—	—
9. *a fi/a W* ♂	—	—	—	—	—	55	49	63	35	202	—	—	—	—	—
10. *a fi/a*^{t} *fi Sd W* ♂	81	58	68	74	281	—	—	—	—	—	69	70	57	85	281
11. *a fi/a Sd W* ♂	—	—	—	—	—	27	25	33	14	99	16	36	23	24	99
12. *a fi/a Sd W* ♀	—	—	—	—	—	28	14	15	14	71	27	15	13	16	71
13. *a fi/a*^{t} *Sd W* ♀	9	5	16	10	40	8	6	15	11	40	6	8	15	11	40
14. *a fi/a*^{t} *Sd W* ♀	10	10	14	14	48	11	9	11	17	48	13	7	11	17	48
15. *a fi/a Sd W* ♀	—	—	—	—	—	12	2	12	16	42	10	4	11	17	42
16. *a fi/a Sd W* ♀	—	—	—	—	—	11	13	10	6	40	13	11	11	5	40
						(*b*) F_2RC									
17. *a fi W/a*^{t} *Sd* ♂	29	27	25	28	109	35	21	26	27	109	36	20	28	25	109
18.* *a fi W/a*^{t} *Sd* ♂	16	15	10	9	50	9	9	8	3	29	11	20	9	10	50
19. *a*^{t} *fi W/a Sd* ♂	27	32	42	42	143	38	21	37	47	143	27	32	35	49	143
20.* *a*^{t} *fi W/a Sd* ♂	19	19	26	22	86	9	5	5	9	28	21	17	22	26	86
						(*c*) F_3RD									
21. *a fi W/a Sd* ♂	—	—	—	—	—	11	8	14	13	46	9	10	15	12	46
22.* *a W/a*^{t} *fi Sd* ♂	22	36	28	32	118	25	15	28	8	76	24	34	21	39	118
23. *a fi W/a*^{t} *fi Sd* ♂	26	29	41	45	141	—	—	—	—	—	22	33	37	49	141
24. *a*^{t} *fi W/a fi Sd* ♂	40	35	50	57	182	—	—	—	—	—	34	41	54	53	182
						(*d*) F_2CD									
25. *a fi/a*^{t} *Sd W* ♂	17	13	19	13	62	12	18	20	12	62	11	19	14	18	62
26.† *a/a*^{t} *fi Sd W* ♂	18	11	12	9	50	—	—	—	—	—	12	17	13	8	50
27. *a fi/a fi Sd W* ♂	—	—	—	—	—	—	—	—	—	—	12	17	14	15	58
28. *a*^{t} *fi/a fi Sd W* ♂	27	30	17	26	100	—	—	—	—	—	22	35	15	28	100
29. *a fi/a W* ♂	—	—	—	—	—	25	28	25	16	94	—	—	—	—	—
30. *a fi/a W* ♂	—	—	—	—	—	19	18	25	15	77	—	—	—	—	—

* Some of these animal's mates were not *fi fi* and their *W-fi* output is omitted; hence, the totals for *W-Sd*, *W-fi* and *W-at* are different.

† None of this animal's mates were *fi fi* and its *W-fi* output is omitted.

Table 10. Single-factor segregations from 'H''s descendants

	W		*a*^{t}		*fi*		*Sd*	
	W:+	χ^2 (1:1)	*a*^{t}:*a*	χ^2 (1:1)	+:*fi*	χ^2 (1:1)	*Sd*:+	χ^2 (1:1)
				(*a*) F_1CC				
total χ^2	—	20·0573	—	5·0523	—	24·7411	—	17·4202
deviation χ^2	848·896	1·3211	274·263	0·2253	780·643	13·1897	692·730	1·0155
heterogeneity χ^2 (d.f.)	—	18·7512 (15)	—	4·8289 (3)	—	11·6600 (14)	—	16·4162 (13)
probability	—	*0·3 to 0·2*	—	*0·2 to 0·1*	—	*0·7 to 0·5*	—	*0·3 to 0·2*
				(*b*) F_2RC				
total χ^2	—	8·4960	—	0·4500	—	2·7553	—	7·8364
deviation χ^2	184·204	1·0309	194·194	0·0000	167·142	2·0227	189·199	0·2577
heterogeneity χ^2 (d.f.)	—	7·4853 (3)	—	0·4500 (3)	—	0·7374 (3)	—	7·5840 (3)
probability	—	*0·1 to 0·5*	—	*0·95 to 0·90*	—	*0·9 to 0·8*	—	*0·1 to 0·05*
				(*c*) F_3RD				
total χ^2	—	13·8672	—	3·1153	—	12·1899	—	10·6807
deviation χ^2	217·280	10·9425	207·234	1·6531	78·44	9·4754	216·271	6·2115
heterogeneity χ^2 (d.f.)	—	2·9920 (3)	—	1·4678 (2)	—	2·9431 (1)	—	4·5269 (3)
probability	—	*0·5 to 0·3*	—	*0·5 to 0·3*	—	*0·1 to 0·05*	—	*0·3 to 0·2*
				(*d*) F_2CD				
total χ^2	—	4·9533	—	5·0529	—	2·0189	—	9·7033
deviation χ^2	235·206	1·9070	110·102	0·3019	126·107	1·5494	113·157	7·1704
heterogeneity χ^2 (d.f.)	—	3·0594 (5)	—	4·7577 (2)	—	0·4726 (2)	—	2·6020 (3)
probability	—	*0·7 to 0·5*	—	*0·1 to 0·05*	—	*0·8 to 0·7*	—	*0·5 to 0·3*

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

TABLE 11. TEST FOR SINGLE-FACTOR HETEROGENEITY BETWEEN GROUPS

	W			a^t		*fi*			*Sd*		
group	*W*:+		χ^2 (d.f.)	a^t:*a*		χ^2 (d.f.)	+:*fi*		χ^2 (d.f.)	*Sd*:+		χ^2 (d.f.)
F_1CC	848	896	1·3211 (1)	274	263	0·2253 (1)	780	643	13·1897 (1)	692	730	1·0155 (1)
F_2RC	184	204	1·0309 (1)	194	194	0·0000 (1)	167	142	2·0227 (1)	189	199	0·2577 (1)
F_3RD	207	280	10·9425 (1)	207	234	1·6531 (1)	78	44	9·4754 (1)	216	271	6·2215 (1)
F_2CD	235	206	1·9070 (1)	110	102	0·3019 (1)	126	107	1·5494 (1)	113	157	7·1704 (1)
total χ^2	—		15·2015 (4)	—		2·1803 (4)	—		26·2372 (4)	—		14·6551 (4)
deviation χ^2	1474	1586	4·0993 (1)	785	793	0·0406 (1)	1151	936	22·1490 (1)	1210	1357	8·4180 (1)
(d.f.)	—		11·1022 (3)	—		2·1397 (3)	—		4·0882 (3)	—		6·2371 (3)
corrected heterogeneity χ^2	—		11·1166	—		2·1397	—		4·1319	—		6·2577
probability	—		*0·02 to 0·01*	—		*0·7 to 0·5*	—		*0·3 to 0·2*	—		*0·1 to 0·05*
probability when F_3RD excluded	—		*0·2 to 0·1*	—		*0·95 to 0·90*	—		*0·98 to 0·95*	—		*0·2 to 0·1*

TABLE 14. SIMULTANEOUS SEGREGATIONS FROM THE SEVEN 'AFFINITY' HETEROZYGOTES

tested animal's probable centrotype			modes of gamete-formation—chromosome V									
			(*P*)		(a^t)		(*Sd*)		(*fi*)		totals	grand totals
			Sd	a^t *fi*	a^t *Sd*	*fi*	a^t *fi* *Sd*	*a*	*fi* *Sd*	a^t		
(I) *W-fi* and *W-Sd* segregating in repulsion												
'H' ♂	$\frac{W\beta\ a\ +\ \alpha\ Sd}{+\alpha\ a^t\ fi\ \beta\ +}$	*W*	45	51	26	31	11	16	0	1	181	340
		+	55	40	26	11	6	16	2	3	159	
			a^t *Sd*	*fi*	*Sd*	a^t *fi*	*fi* *Sd*	a^t	a^t *fi* *Sd*	*a*		
25 ♂	$\frac{W\beta\ a^t\ +\ \alpha\ Sd}{+\alpha\ a\ fi\ \beta\ +}$	*W*	9	11	1	6	1	2	0	0	30	62
		+	11	9	3	3	0	5	0	1	32	
											grand total	402
(II) *W-fi* and *W-Sd* segregating in coupling												
			Sd	*fi*			*fi* *Sd*	*a*				
4 ♂	$\frac{W\alpha\ a\ +\ \alpha\ Sd}{+\beta\ a\ fi\ \beta\ +}$	*W*	22	4	—	—	5	4	—	—	35	82
		+	22	4	—	—	5	4	—	—	35	
7 ♂	$\frac{W\alpha\ a\ +\ \alpha\ Sd}{+\beta\ a\ fi\ \beta\ +}$	*W*	19	10	—	—	6	4	—	—	39	89
		+	17	25	—	—	0	8	—	—	50	
15 ♀	$\frac{W\alpha\ a\ +\ \alpha\ Sd}{+\beta\ a\ fi\ \beta\ +}$	*W*	9	1	—	—	1	3	—	—	14	42
		+	9	14	—	—	2	3	—	—	28	
			a^t *Sd*	*fi*	*Sd*	a^t *fi*	*fi* *Sd*	a^t	a^t *fi* *Sd*	*a*		
14 ♀	$\frac{W\alpha\ a^t\ +\ \alpha\ Sd}{+\beta\ a\ fi\ \beta\ +}$	*W*	7	3	3	3	3	—	—	1	20	48
		+	5	6	2	7	4	2	—	2	28	
			Sd	a^t *fi*	a^t *Sd*	*fi*	a^t *fi* *Sd*	*a*	*fi* *Sd*	a^t		
19 ♂	$\frac{W\alpha\ a^t\ fi\ \beta\ +}{+\beta\ a\ +\ \alpha\ Sd}$	*W*	17	10	5	6	4	8	1	8	59	143
or	$\frac{W\alpha\ a\ +\ \alpha\ Sd}{+\beta\ a^t\ fi\ \beta\ +}$	+	19	26	8	13	7	9	1	1	84	
											grand total	404

TABLE 19. SUMMARY OF STATISTICAL EVIDENCE

bodies of data	description of data and significance tests		χ^2 (d.f.)	probability: A (low: see key)	probability: B (high: see key)
initial observations	(1) 1950 Non-random segregation of *Ca*, *W*, and *fi*; *W-fi* r.f. = 40 % C				
	Test of triple independence:		9·7568 (1)	0·01 > *p* > 0·001	—
	(2) 1951 Non-random segregation of *W* with *fi* and *Sd*, and independence of *W* with a^t, *W-fi* r.f. = 42 %, *W-Sd* = 45 %, *W-*a^t = 50 % (all C and R)				
	Test of independence:	*W-fi*	13·7813 (1)	0·001 > *p*	
		W-Sd	9·0134 (1)	0·01 > *p* > 0·001	—
W-V investigation	(3) 1952 Selection of male progenitor 'H'. Probability of obtaining, out of seven trials, one animal with a χ^2 testing independence *W*/*fi* as high as 6·8681 is:		—	*p* = 1 in 16	—
	'H''s performance (91 progeny) consistent with expectation, for deviation of *W-*a^t from 50 % < that of *W-Sd* < that of *W-fi*. *W-fi* and *W-Sd* r.f.'s significantly > 50 %, a possibility expected only if affinity operating. Test of homogeneity of 'H''s later performance (249 progeny) with earlier (91 progeny):		2·4607 (1)	—	0·2 > *p* > 0·1
	(4) 1955 Performance of 'H''s descendants (thirty animals tested; some 2500 progeny bred)				
	(i) Fluctuating single-factor disturbance as possible cause of associations is obviated by omission from further analysis of the four animals causing single-factor heterogeneity				
	(ii) Consistent viability disturbance, as possible cause of associations, is disproved by heterogeneity between pairs of complementary genotypes				
	Test of homogeneity:	*W-fi*	41·5019 (21)*	0·01 > *p* > 0·001	—
		W-Sd	31·7284 (21)*	0·1 > *p* > 0·5	—
	(iii) Linkage, as possible cause of association, is disproved by insignificance of inequality of non-recombinant pair and recombinant pair (on total progeny from twenty-six animals)				
	Test of equality:	*W-fi*	0·8555 (1)	—	0·5 > *p* > 0·3
		W-Sd	2·6327 (1)	—	0·2 > *p* > 0·1
	(iv) Independence (and linkage) is disproved by significant heterogeneity between non-recombinant pair and recombinant pair (as between twenty-six animals)				
	Test of homogeneity:	*W-fi*	41·8601 (21)*	0·01 > *p* > 0·001	—
		W-Sd	35·7351 (21)*	0·05 > *p* > 0·02	—
	(v) Affinity, as possible cause of association, is supported both by significant heterogeneity between complementary pairs, and by significant excess of pair expected to be greater (when quasi-linkage occurs) over pair expected to be less (on total progeny from twenty-six animals)				
	Test of equality:	*W-fi*	3·5059 (1)	0·1 > *p* > 0·05	—
		W-Sd	7·8769 (1)	0·01 > *p* > 0·001	—
	Test of homogeneity:	*W-fi*	39·2630 (21)*	0·01 > *p* > 0·001	—
		W-Sd	30·56023 (21)*	0·1 > *p* > 0·05	—
	N.B. One animal's performance at first appeared to show a deviation in a direction expected rarely on the basis of affinity. Its parents were then found to be, by error, an intercross for *W*; from such a mating the expected frequency of an animal showing a deviation in the direction observed is:		—	—	*p* = 0·1641
	and in the opposite direction is:		—	0·02 > *p*	—
	This animal's performance is therefore entered into the table above as showing a deviation in the expected direction				
	(vi) A group by group analysis also supports the hypothesis of affinity (see text)				
	(vii) Affinity is further supported by agreement between observed number (5) of animals showing quasi-linkage with number expected (6·36) out of 26 tested. (5 % level used for discrimination):				
	Test of agreement:		0·3850 (1)	—	0·7 > *p* > 0·5

Key: C = Coupling backcrosses, R = repulsion backcrosses, r.f. = recombination fraction.

A. These are the cases in which a low probability supports the general argument that a new phenomenon is operating and that it is probably affinity. B. These are the cases in which a high probability supports the argument.

* There are only 21 d.f., not 25, because, although the twenty-six animals are included in the analysis, four segregated in *fi* and not *Sd* and four in *Sd* and not *fi*.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES