PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 5 FEBRUARY 15, 1919 Number 2

IS THE ARRANGEMENT OF THE GENES IN THE CHROMOSOME LINEAR?

By W. E. Castle
Bussey Institution, Harvard University

Read before the Academy, November 18, 1918

Every biologist is familiar with the remarkable discoveries of Morgan and his associates concerning the germ-cells of Drosophila. One of the most important of these discoveries is concerned with the phenomenon of linked inheritance. This kind of inheritance, while entirely conformable with Mendel's law, forms a very distinct and important class of cases whose existence has been brought to light since the rediscovery of the general law in 1900. Under the general law it is found that characters which behave as distinct units in heredity assort quite independently of each other. Thus if parents are crossed one of which possesses two characters, A and B, while the other lacks them, then the offspring of this cross will transmit A and B sometimes associated in the same gamete, sometimes in different gametes, the two events being under the laws of chance equally probable.

But in linked inheritance, a phenomenon first made known to us through the work of Bateson and his associates in England, later more fully explored and explained by Morgan, A and B are not wholly independent of each other in transmission. If they enter a cross together, they have a tendency to stay together in later generations; and if they enter a cross separately, they have a tendency to remain apart in later generations. Morgan has suggested that what binds or links two characters together is the fact that their genes lie in the same body within the cell-nucleus. Such bodies he supposes are the chromosomes. The evidence for this conclusion is very strong. Morgan and his associates have demonstrated the existence in Drosophila of four groups of linked genes corresponding with the four pairs of chromosomes which the cell-nucleus of Drosophila contains. Morgan has further suggested (and has beyond doubt established the fact) that the genes within a linkage system have a very definite and constant relation to each other. He supposes their

arrangement to be linear and in the group of genes most exhaustively studied, that of the 'sex chromosome' has represented them in a 'chromosome map, as shown in Diagram I.

That the arrangement of the genes within a linkage system is strictly linear seems for a variety of reasons doubtful. It is doubtful, for example, whether an elaborate organic molecule ever has a simple string-like form. Let us, therefore, examine briefly the evidence for or against the idea of linear arrangement of the genes. It is supposed by Morgan that two genes lying in the same chromosome show close linkage if they lie close together, but less linkage if they lie farther apart, and that the farther apart they are the less will be their linkage. As a measure of the distance apart of two genes he takes the percentage of cross-overs between them. This term requires a word of explanation. If two genes, A and B, enter a cross in the same gamete and emerge from it in different gametes, one of them has evidently *crossed-over* from the chromosome in which it lay to the mate of that chromosome (all chromosomes being paired prior to the formation of gametes). Likewise if the two genes, A and B, having entered a cross separately (being contributed by different parents), later emerge from the cross together, it is evident that one of them has again *crossed-over* so as to lie in the same chromosome as the other. The readiness with which cross-overs occur between two genes will on Morgan's hypothesis depend on their distance apart and the percentage of cross-overs between genes will be proportional to the distances between them. These assumptions have abundantly proved their utility as a working hypothesis, for it has been found possible, knowing what certain cross-over values are, to predict others with a fairly good degree of accuracy.

If the arrangement of the genes is strictly linear, so that A, B, C, etc., lie in a straight line, then it should be possible to infer the distance AC, if the distances AB and BC are known, since AC = AB + BC. But if the distance AC is less than the sum of AB and BC, then the arrangement can not be linear, since B will lie out of line with A and C. In reality it has been found that the distances experimentally determined between genes remote from each other are in general less than the distances calculated by summation of supposedly intermediate distances, and the discrepancy increases with increase in the number of known intermediate genes. To account for this discrepancy Morgan has adopted certain subsidiary hypotheses, of 'interference,' 'double crossing over,' etc., in accordance with which it is supposed that cross-overs between nearer genes interfere with or lessen the apparent amount of crossing-over with genes more remote. He therefore bases his chromosome map on summation of the shorter distances. This, however, leads to results which can be shown to be impossible.

Morgan's map of the sex-chromosomes places five out of twenty-nine genes at distances between 55 and 66 from the zero end of the chromosome, where yellow is located. A moment's reflection will show these to be impossible re-

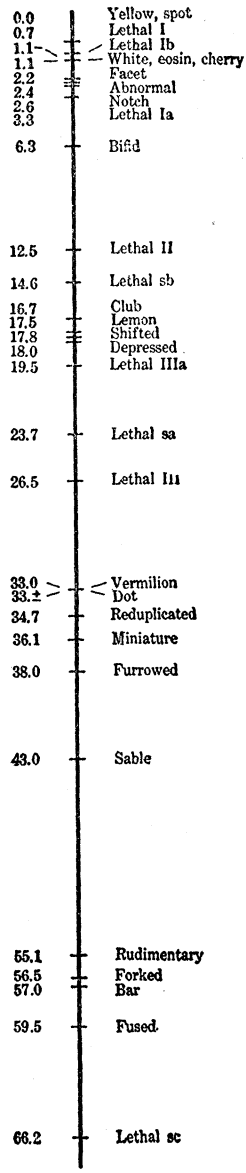

DIAGRAM I.

"Diagram I shows the relative positions of the genes of the sex-linked characters of *Drosophila*. One unit of distance corresponds to 1% of crossing-over."

Morgan and Bridges, p. 23.

lations, for a cross-over percentage greater than fifty is absurd. If A and B assort wholly independently, without any linkage whatever, just as they would in ordinary Mendelian inheritance where no linkage exists, cross-overs and non-cross-overs will be equal, 50% each. Any cross-over value consistently less than 50% shows linkage. *A cross-over value greater than fifty cannot exist.* For there must be either linkage or no-linkage. But no-linkage means 50% cross-overs, and linkage means less than 50% cross-overs. Hence a value greater than 50% cannot occur.[1] As an alternative to the hypothesis of linear arrangement it is possible to suppose that the arrangement of the genes is not linear, that the nearer genes are not directly in line with the more remote ones.

If the arrangement of the genes is not linear, what then is its character? This query led me to attempt graphic presentation of the relationships indicated by the data of Morgan and Bridges[2] but finding this method unsatisfactory I resorted to reconstruction in three dimensions, which has proved very satisfactory. The data used consist of the experimentally determined cross-over percentages between twenty genes of the sex-chromosome of Drosophila, as given in Table 65 of Morgan and Bridges. The only hypothesis involved in the reconstruction is Morgan's fundamental one that distances between genes are proportional to cross-over percentages. The secondary hypothesis, that distant genes are really farther apart than indicated by the experimental data, is rejected for the reason already explained, that impossible relationship are thereby entailed. Taking the data, then, exactly as they stand, we find it possible to make a very complete and on the whole self-consistent reconstruction of the architecture of the sex-chromosome from the linkage relations of its genes. A small ring of wire is taken to represent the locus of a gene. Two genes are placed as far apart (in inches) as there are units in the cross-over percentage between them. Thus between yellow-body and white-eye there is shown by Morgan's data to be a cross-over value of 1.1%. Consequently the rings which represent these genes in the reconstruction are joined by a wire 1.1 inches long. Between white and vermilion the cross-over percentage is 30.5. Accordingly the connecting wire in this case is made 30.5 inches long. Proceeding in this way the reconstruction shown in figures 1 and 2 is obtained. It indicates an arrangement of the genes not by any means linear, but rather in the form of a roughly crescentic plate longer than it is wide, and wider than it is thick. It is shown in side view in figure 1, and in edge view in figure 2.

That the arrangement of the genes can not by any possibility be linear is shown by two critical cases, the loci for *bifid* and *abnormal*. Bifid (*Bi*, figs. 1 and 2) is shown by a series of over 3,600 observations to be at a distance 5.5 from yellow. It is almost the same distance from white, viz., 5.3, as shown by 23,595 observations. Yet white and yellow are distant from each other only 1.1 units, as shown by 81,299 observations. Therefore bifid can lie neither above nor below yellow and white, in the line which joins them, but

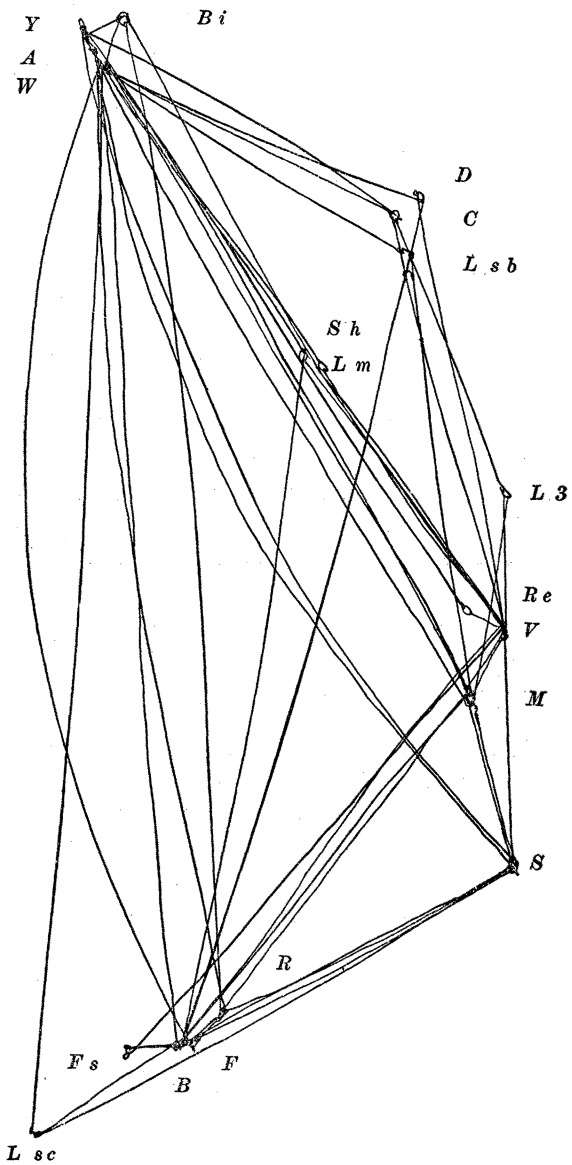

FIG. 1. SIDE VIEW OF MODEL

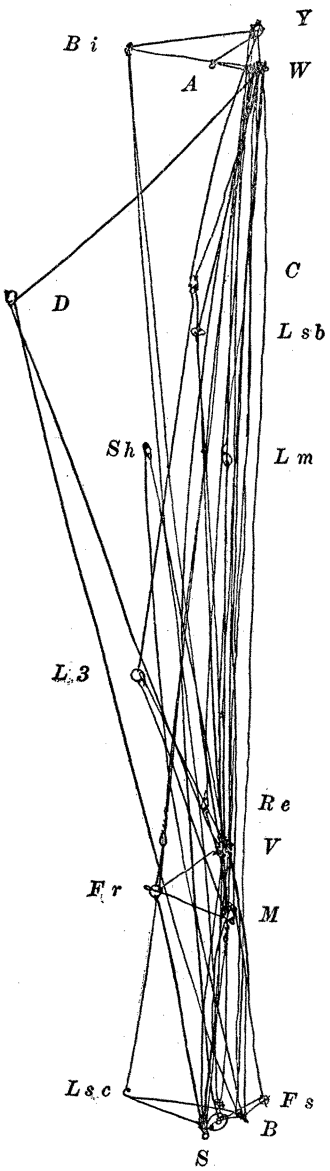

FIG. 2. EDGE VIEW OF MODEL

Seen at right angles to figure 1.

Showing relative positions of genes of 20 sex-linked characters of Drosophila, linear arrangement not being assumed. For significance of letters, compare Diagram I.

must lie laterally about equidistant from both. Other linkage relationships of bifid show that it lies as far as possible removed from the plane in which most of the genes lie. The other case mentioned is that of abnormal abdomen (*A*, figs. 1 and 2). Its linkage relations are known with only two other genes, yellow and white. But the relation of these two with each other is one of the best determined and the linkage of abnormal with each of them rests on more than 15,000 observations in each case. The yellow-white linkage is 1.1, as already stated; abnormal-yellow is 2.0, and abnormal-white 1.7. These relations make it impossible for abnormal to lie in line with yellow and white. Until a third linkage relation of abnormal is determined, it may swing freely round the line which joins yellow with white but can never come into line with them. A third linkage relation having once been determined for abnormal, its linkage with any other gene in the sex-chromosome could be readily predicted from direct measurement of the reconstructed figure. The actual test of the utility of this method of portraying linkage relationships could easily be made by first forecasting by measurement what undetermined linkage values are likely to be and then actually determining them by experiment. Such predictions could not fail to come nearer the truth than predictions based on a linear map, if as I have suggested the arrangement is really not linear.

What, it might be asked, does this reconstruction signify? Does it show the actual shape of the chromosome, or at any rate of that part of it in which the observed genetic variations lie? Or is it only a symbolical representation of molecular forces? These questions we can not at present answer. A first step toward answering them will be the construction of a model which will give us reliable information as to undetermined genetic relationships. A model which will answer questions truthfully must be a truthful presentation of actual relationships even though we do not know whether they are spatial or dynamic.

If the arrangement of the genes in the chromosome is not linear, Morgan's theory of linkage must be somewhat modified. (1) The fundamental assumption that the genes lie in the chromosomes and have a definite orderly arrangement there is not disturbed. (2) The further assumption that the respective distances between the genes determine their closeness of linkage one with another may also stand unchallenged. (3) But the assumption that the arrangement of the genes within the chromosome is linear cannot be accepted without proof, which at present is lacking. This assumption has made necessary other secondary assumptions, likewise unproved, which are superfluous if this one is abandoned.

Such an unproved secondary hypothesis is that of double-crossing-over. The experimental data show that double-crossing-over *must occur*, if the arrangement of the genes is linear. For if three genes, A, B, C, are linear in their arrangement in the order named, and all lie in the same gamete, and if subsequently A and C are found together in one gamete and B in another,

it is evident that this rearrangement can have come about only as a result of two breaks in the linkage chain, viz., one between A and B, and another between B and C. But if the arrangement is *not* linear, double-crossing-over need not be assumed as an explanation of the observed regroupings. For if A, B, and C are linked in a triangle, not in a straight line, then B may be freed from its connections with A and C without necessarily disturbing the connection of A and C with each other. Freeing of B will involve no greater number of breaks than the freeing of either A or C. It will still be true, however, as indicated by the experimental data, that certain groupings of three particular genes are easier to obtain than others. Thus in the case of the three genes white, bifid and vermilion, it is hardest to obtain the regrouping which involves detaching bifid from the other two. Morgan assumes that this is because bifid lies *between* the other two in a single linkage chain and so could be detached only by two breaks; it is possible, however, that the reason may be that bifid lies peripherally in the linkage system and could be detached only by an oblique longitudinal break, whereas either of the others could be detached by a simple transverse break. Similarly in the trio, white-vermilion-sable, it is vermilion which is difficult to detach; and in the group, vermilion-sable-bar, it is sable. Always it is the *middle one* considered with reference to the long axis of the system. This may be because, as Morgan supposes, only transverse breaks occur, of which two taking place simultaneously are required to produce the difficult regrouping, or it may be because transverse breaks are more frequent than oblique longitudinal ones, of which a single one would suffice to accomplish the regrouping, if the genes are not strictly linear in arrangement.

The phenomenon of 'coincidence' as described by Weinstein[8] is this. If crossing-over occurs toward one end of a chromosome, it is less likely to occur simultaneously elsewhere in the same chromosome. Crossing-over in one part of a chromosome is thus supposed to 'interfere with' crossing-over elsewhere in the same chromosome. If we adopt the hypothesis of linear arrangement, interference must be assumed to occur. Observed facts require this. But if we do not adopt this hypothesis but suppose that what have been called 'double cross-overs' are really the result of single oblique or single longitudinal breaks, then the supposed phenomenon of interference may mean only this, that transverse breaks are more likely to occur than longitudinal ones.

Finally, if the genes are not arranged in a single linear chain, the chiasma-type theory will need to be reëxamined. Such a purely mechanical theory seems inadequate to account for interchange of equivalent parts between twin organic molecules, such as the duplex linkage systems of a germ-cell at the reduction division must be. It seems more probable that preceding the reduction division a period of instability within the chromosome molecule comes on. Twin molecules are now closely approximated and parts of one may leave their former connections and acquire new connections with the

corresponding parts of the other twin. It is evident from the experimental data, notably that of Muller,[4] that new connections are not formed with any torn fragment of chromosome which happens to come into the proper position, but that connections are always formed at exactly corresponding points with homologous systems of genes.[5] It is like the replacement of one chemical radicle with another within a complex organic molecule and it seems highly probable that such is its real nature.

[1] The distances shown in Morgan's chromosome map in excess of 50 (admittedly not obtained experimentally but only by summation) are therefore too large. Accordingly, if one clings to the assumption that the arrangement of the genes is linear, it must be, not that the longer distances are too short, as Morgan has assumed, but that the short distances are too long. Therefore, any hypotheses framed to account for an apparent shortening of the long distances are superfluous. The long distances given by direct experiment are long enough; they approach the limit of the possible, viz., 50%. Thus in table 65 of Morgan and Bridges, we find the following high cross-over percentages given by direct experiment:—yellow-bar, 47.9; white forked, 45.7; and white-lethal *sc*, 46.0. What is needed therefore, if the linear arrangement hypothesis is retained, is a secondary hypothesis to explain why the short distances given by experiment are too long.

But if we abandon the hypothesis of linear arrangement, all secondary hypotheses are unnecessary. The experimentally obtained cross-over percentages may be accepted at their face value, which in every case fall within the limits of the possible, 0 and 50.

[2] Morgan, T. H., and Bridges, C. B., Sex-linked inheritance in Drosophila. *Carnegie Inst. Washington, Publ.*, No. 237, 1916, (88 pp., 2 pl.).

[3] Weinstein, A., *Genetics*, **3**, 1918, (135–172).

[4] Muller, H. J., *Amer. Nat., Lancaster, Pa.*, **50**, 1916.

[5] The case of 'deficiency' studied by Bridges (*Genetics*, **2**, pp. 445–460, Sept. 1917) forms an apparent exception to the rule. Here a certain segment of the linkage system was as regularly wanting as it is commonly present. The regularity of the process, however, shows that the principle of union at particular points still holds. In the deficiency race, a new and simplified linkage system had been established and this persisted.