

THE SPATIAL RELATIONS OF GENES

BY A. H. STURTEVANT, C. B. BRIDGES, AND T. H. MORGAN

COLUMBIA UNIVERSITY AND CARNEGIE INSTITUTION OF WASHINGTON

Communicated April 11, 1919

Castle¹ has proposed an arrangement of linked genes in three dimensions, based on the assumption that the distance between any two loci is exactly proportional to the observed crossover value. He believes that this system gives a better agreement between observation and expectation than does the hypothesis of a strictly linear arrangement that we have developed.

According to Castle, "that the arrangement of the genes within a linkage system is strictly linear seems for a variety of reasons doubtful." The only one of these reasons specified is that "it is doubtful, for example, whether an elaborate organic molecule ever has a simple string-like form." The suggestion that chromosomes are organic molecules is probably not intended to be taken seriously. The further argument, that organic molecules probably never have a simple string-like form, is scarcely to the point, for *chromosomes* do have a thread-like form. The comparison between crossing over and organic substitution reactions seems forced, for these reasons, as well as on strictly chemical grounds.

Castle presents two cases as "critical" ones showing that the arrangement cannot be linear; viz., the relations of bifid and of abnormal to yellow and white. The case of abnormal presents some complications,² but bifid may be accepted as a crucial case. On the basis of the data summarized by Morgan and Bridges³ and used by Castle, the crossover values for yellow, white and bifid are:

| | |
|-------------------|-----|
| Yellow white..... | 1.1 |
| White bifid..... | 5.3 |
| Yellow bifid..... | 5.5 |

Each of these values is based on several thousand flies, so that the probable errors arising from random sampling are small. Castle points out that no one of the three values is either the sum or the difference of the other two, and therefore concludes that the placing of the three loci in a line cannot represent their relations to each other. But these data come from several sources; and it has frequently been pointed out that crossover values are subject to variation, due to genetic factors,⁴ to environmental causes,⁵ or to differential viability.⁶ In such a case as the present one, then, strictly comparable data can be obtained only from experiments in which all three loci are followed at the same time.⁷ Such crucial data for yellow, white, and bifid had been given by Morgan and Bridges³ and by Muller,⁴ in papers which were known to Castle. These experiments may be summarized as follows:

| | NON-CROSS-OVERS | YELLOW SEPARATES FROM OTHER TWO | WHITE SEPARATES | BIFID SEPARATES | TOTAL |
|-------------------------|-----------------|---------------------------------|-----------------|-----------------|-------|
| Morgan and Bridges..... | 487 | 3 | 0 | 16 | 506 |
| Muller..... | 673 | 12 | 0 | 27 | 712 |
| Total..... | 1160 | 15 | 0 | 43 | 1218 |
| Percentage..... | | 1.2 | 0.0 | 3.5 | |

Here we have crossover values that are strictly comparable with each other, and they show:

| | |
|-------------------|-----|
| Yellow white..... | 1.2 |
| White bifid..... | 3.5 |
| Yellow bifid..... | 4.7 |

These relations, in which the sum of the smaller values is exactly equal to the larger one, can be represented⁸ only by placing the three loci in a straight line, in the order yellow, white, bifid.⁹

It will be seen that this linear arrangement is necessitated by the fact that white never separated from yellow and bifid; i.e., no 'double crossover' occurred. By considering data involving only short distances (e.g., white bifid club, bifid club vermilion, club vermilion miniature, vermilion miniature sable, sable rudimentary forked, rudimentary forked bar, and forked bar fused; or still better, certain unpublished data including the loci ruby, cut, tan, garnet, etc.), it is possible to handle the whole X-chromosome of *Drosophila* as made up of successive overlapping sections in which double crossing over either does not occur or is so rare as to be negligible. Since these sections must be represented as straight lines, and since they overlap, the whole X-chromosome must be represented as a straight line.

For longer sections of this same chromosome, however, double crossovers do occur, so that the distances apart on the straight line are no longer proportional to the crossover values. The 'map-distance' for white forked, for example, is 55.4, though the crossover value observed¹⁰ is 43.9.

Castle is disturbed by such map distances greater than 50. He says: "A crossover value greater than fifty cannot exist. For there must be either linkage or no-linkage. But no-linkage means 50% crossovers, and linkage means less than 50% crossovers. Hence a value greater than 50% cannot occur." It is evident that the conclusion of this curious syllogism depends solely on the definition of linkage contained in the second half of the second premise. It is true that crossover values significantly greater than fifty have never been found, but this is due to the frequency of double crossovers—to the fact that 'coincidence' is high for long distances. But coincidence is known to be a variable quantity, so that there would seem to be no *a priori* reason to suppose that crossover values greater than fifty are impossible. Furthermore, as has been explained above, map distances greater than 50 have never been intended to represent observed crossover values greater than fifty.

It is to be observed that the three dimensional figure given by Castle might seem capable of reduction to a curved line lying in one plane. An examination of his figure 2 shows that the only loci far from the plane that includes most of the group are bifid, depressed, shifted, lethal 3, furrowed, and lethal sc. Let us then examine these individually.

Bifid has already been shown to lie in the same straight line with yellow and white, both of which are in Castle's thickly populated plane.

Morgan and Bridges give three crossover values for depressed, two of which (white depressed and depressed vermilion) are stated to be based on 59

flies each. Such meager data are clearly insufficient to establish accurate crossover values. But the values given by Morgan and Bridges are themselves incorrect, as is shown by their account of the actual experiment (p. 68). The total number of flies is 69, not 59, and the crossover values accordingly become 17.4 for white depressed (instead of 20.3) and 14.5 for depressed vermillion (instead of 17.0). The sum of these values, 31.9, is not significantly different from 30.5, the accepted value for white vermillion. This relation, and the fact that the 69 flies included no double crossover, give no warrant for supposing that the locus of depressed is outside of the plane that includes white, vermillion, and bar.

Only two crossover values are given for shifted—shifted vermillion and shifted bar. It is evidently impossible to place a point in space by means of its distances from only two other points. Three are necessary, as Castle himself points out in connection with abnormal and bifid. There is thus no reason whatever for placing shifted out of the common plane.

The aberrant position of lethal 3 in Castle's model is evidently due in large part to the fact, shown by Morgan and Bridges (p. 75), that the experiments involving this locus gave a value of 2.3 for vermillion miniature, which is less than the mean value, 3.1, used by Castle. The values obtained in the lethal 3 experiments were:

| | |
|---------------------------|-----|
| Lethal 3 vermillion..... | 6.8 |
| Vermillion miniature..... | 2.3 |
| Lethal 3 miniature..... | 9.3 |

The third value is substantially the same as the sum of the first two—hence the three loci must be represented as lying very nearly in a straight line.

In the case of furrowed, the values are all based on a few flies (240 or less in every case), and the three most important values (vermillion furrowed, miniature furrowed, and furrowed sable) were all obtained in different experiments, so that comparisons are hazardous. Furrowed is a character that is often difficult and sometimes impossible to classify, so that there is a large probability of error in the counts that have been reported.

In the case of lethal *sc* there is only one available crossover value (bar lethal *sc*, 8.3) small enough so that double crossing over would be negligible; and there are no data involving lethal *sc* and more than one other locus at the same time. The data are therefore not of the type to give a decisive answer as to the relations of the locus.

A careful examination of the data thus shows that a single plane suffices for the representation of the loci dealt with by Castle. Within this plane the positions of the remaining loci fall, in Castle's figure, roughly into a single curved line. The only noticeable exception is the locus of lemon. This locus is based on only 241 flies, and these gave rather too small a crossover value for the well established white vermillion distance, so that this discrepancy is not significant. The arrangement of the genes is thus approximately linear, but

the line is curved instead of straight. This curvature is due to the phenomenon of double crossing over, and the degree of curvature is dependent on the frequency of double crossovers; i.e., upon 'coincidence.' This method of representing coincidence, however, leads to inconsistencies. As was pointed out above, it is possible to build up the whole X-chromosome from successive overlapping sections, each of which is so short as to include few or no double crossovers. Yet double crossovers are not rare in longer sections. If ABC and BCD are sections so short that no double crossing over occurs within each, section ABCD may be long enough so that double crossing over in which one crossover is between A and B and the other between C and D may still occur; and in that case it becomes necessary to represent D in two positions at the same time. Such situations are actually known in *Drosophila*. Hence the curved line cannot give a consistent scheme; nor can any other scheme based on the assumption that long distances, as well as short ones, are to be represented as proportional to observed percentages of crossing over.

According to Castle the supposition that double crossovers do occur is "an unproved secondary hypothesis." Of course, if three genes are imagined as not lying in a straight line, a single plane may separate any one from the other two. Under these conditions double crossing over has no meaning. But the fact remains that when three linked genes are studied simultaneously, one pair of contrary classes is *always* small, and this class is always the one that is the double crossover class on the linear arrangement scheme. Some method of accounting for the smallness of this class is evidently demanded. Castle recognizes this, and suggests that the explanation may be either that "only transverse breaks occur, of which two taking place simultaneously are required to produce the difficult regrouping" (i.e., the familiar double crossover explanation), or that "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." Muller⁴ has published data bearing on this point. He followed simultaneously eleven loci in the X-chromosome of *Drosophila*, and obtained some double crossover classes that according to Castle's model are impossible with a single plane. However, if the loci all reduce to the curved line discussed above, than any so-called double crossover is possible with a single separation plane. But those classes now termed triple and quadruple crossovers are all impossible with a single plane of crossing over. In the case of the X-chromosome over a dozen triple crossovers have been observed, and in the case of the second chromosome they are so frequent that 131 were observed in a single experiment. It follows that, even if the genes are not arranged in a straight line, the occurrence of double crossing over is an established fact, not an unproved hypothesis.

If crossover planes do not occur longitudinally, or occur thus less often than transversely, it is difficult to see how the distances apart of the loci in Castle's model can be proportional to the crossover values, except in the case where

the line joining the loci is perpendicular to the average direction of crossover planes. The crossover value for lines in any other directions would seem to be proportional to their *projections* on such a perpendicular. In short, both of Castle's alternative subsidiary hypotheses to account for the rarity of double crossovers are inconsistent with his primary assumption that distances are proportional to crossover values.

¹ These PROCEEDINGS, 5, 1919, (25).

² Abnormal is a character that is separable from wild-type only under certain environmental conditions; and is also simulated by accidental abnormalities. It is, therefore, not a reliable character to use in linkage experiments.

³ *Carnegie Inst., Washington, Pub.*, No. 237, 1916.

⁴ Sturtevant, these PROCEEDINGS, 3, 1917, (555), and Muller, *Amer. Naturalist*, 50, 1916, (193, 284, 350, 421).

⁵ Bridges, *J. Exp. Zool.*, 19, 1915, (1), and Plough, *Ibid.*, 24, 1917, (147).

⁶ Bridges and Sturtevant, *Bio. Bull.*, 26, 1914, (205), and Morgan and Bridges, loc. cit., note 3.

⁷ That Castle tacitly recognized this is shown by the fact that he omitted lethal 2 from his model, without even mentioning the fact. It is obvious from an examination of the data in this case that the crossover values are inconsistent with those obtained in other experiments not involving lethal 2.

⁸ Provided the relations of the loci are to be represented by straight lines joining them and proportional in length to the corresponding crossover values.

⁹ It does not follow that 4.7 is to be taken as the locus of bifid, for the mean value indicated by all bifid experiments is the one most likely to coincide with future experiments; i.e., to give the best predictions.

¹⁰ Castle concluded that the white forked value of 45.7, used by him, is somewhat too high, which is true; but there were available to him more than 40,000 flies (Bridges, *Genetics*, 1, 1916, (1), and Weinstein, *Ibid.*, 3, 1918, (135)) giving the lower value, in contrast to the less than 4000 flies on which the high value was based.
