## THE ORDER OF THE GENES

THE proof of the linear order of the genes is derived directly from the linkage data. It is not dependent on the chromosome theory of heredity. Fortunately, as was pointed out in the last chapter, there are many facts about the maturation stages of the eggs and sperm that fit in extraordinarily well with the theory of the linear order of the genes, but let me repeat, the proof of the order is not dependent on the chromosomal situation. The evidence for the linear order is furnished by linkage and its correlative phenomenon, crossing over. By linkage is meant that certain factors that enter the cross from each parent remain together in subsequent generations, more often than they separate. For example, if in *Drosophila* yellow wings and white eyes have entered from one parent and gray wings and red eyes from the other, the new (crossover) combinations, yellow and red, gray and white, are less numerous than are the original linked combinations, yellow and white, gray and red. The number of individuals (crossovers) that result from this interchange, expressed as percentage of the whole number of individuals, is called the crossover value. Such a percentage indicates how often the linkage is broken. Thus, if crossing over between yellow and white is shown in 1 per cent. of the gametes, then 1 stands for the crossover value of yellow and white. Conversely, yellow and white have remained together (linked) in 99 per cent. of the gametes. We speak of the linkage relations in such cases in terms of the crossover values, here 1 per cent.

For the proof of the linear order of the genes, it is only

necessary to represent one set of linked genes (a, b, c, etc), ignoring the normal allelomorphic series, for these follow the same (reciprocal) changes.

If $a, b,$ and c stand for three genes, and if the linkage relations of a to b and of b to c are known, the relation of a to c is a function of the sum of ab and bc or of the difference of ab and bc. For example, if the crossover value ab is expressed as 5, and that of bc as 10, then ac is a function of the sum (15), or the difference (5) of ab and bc. It cannot be said that ac must be 5 or 15 because another possible process may intervene to affect the sum or the difference, *viz.,* double crossing over in the region involved. By making the distance so small that double crossing over is practically excluded the sum or the difference is actually the realized result, as the following example illustrates:

When three mutant characters yellow, white and bifid were all used together in a single experiment, it was found that there were 1160 non-crossovers, 15 flies representing single crossovers between yellow and white, and 43 flies representing single crossovers between white and bifid. There were no flies representing crossing over in both regions at the same time, *i.e.,* there were no double crossovers. Thus the crossover value yellow white is 1.2, and the crossover value white bifid is 3.5. The same data give the yellow bifid crossover value of 4.7, which is precisely the sum of the two component values:
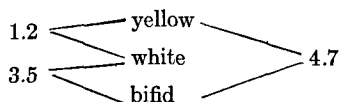
The simplest way in which such a relation can be thought of is that the three genes stand in a line. Suppose a fourth linked gene, d, is added to the series. It is then found that bd, is a function of the sum or of the differ-

ence of b to c and c to d. Four points arranged in a
straight line still fulfill the relations here found.   I know
of no other geometric configuration that covers all these
results—perhaps there is none.   When we add more
and more linked genes to the series, and find the same
predictable relations continue to hold, the theory of
the linear arrangement becomes firmly established.   Per-
haps the best proof of the linear order is found in the
opportunity it gives for prediction; for, when the rela-
tion of d to b and to c is known its relation to a can be
foretold accurately.

It has been found when there is a large amount of cross-
ing over between two factors used in an experiment, that
the crossover value is not the same as the value deter-
mined by adding together the crossover values of inter-
mediate points between the two factors in question.   What
appears here to be a contradiction proves, when under-
stood, to be one of the best pieces of evidence in support
of the theory of the linear order.

A few examples will serve to illustrate the point at
issue.   When a fly with yellow wings and bar eyes is mated
to a wild-type fly, the amount of crossing over in the F_1
female between yellow and bar (as determined by back-
crossing) is 43.6 per cent., but if the crossing over between
yellow and bar is calculated by adding up the crossover
values obtained by using intermediate points $(ab + bc$,
etc.) the value is about 57 per cent.   The apparent incon-
sistency is at once cleared up by arranging the experiment
so that, while obtaining the data for yellow and bar,
there are also obtained data showing what is happening in
the region between them. It is found that a large amount of
double crossing over occurs, and, when the correction for
this is made, the ''discrepancy'' disappears.   If crossing
over may take place at any level, it is obvious that it might
occur at two points at the same time, and experience

shows that such is the case, for such double crossing over can be detected if enough points in the series are "involved" to catch all single crossovers. Now, as shown in Fig. 52, whenever double crossing over takes place between y and B, the two series that result, as marked by their ends alone (y and B), are still y and B. The flies will therefore be placed in such classes, which are the non-crossover classes. A numerical increase in this class will decrease the calculated percentage of crossovers. Thus double crossing over by increasing the number of
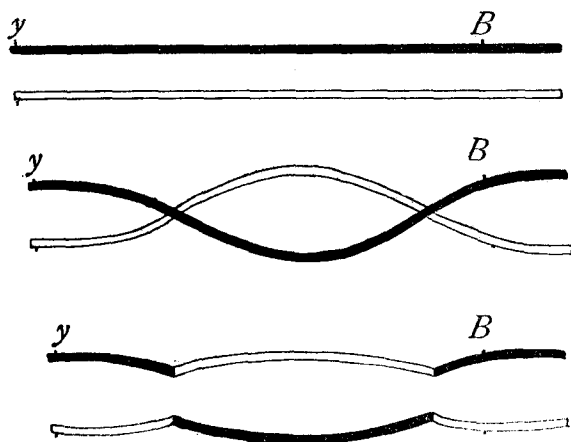

Fig. 52.—Scheme illustrating how double crossing over between two distinct genes, y and B, is not recorded, when only y and B are involved.

apparent non-crossovers, decreases the observed percentage of crossovers. When enough points are marked along the series to pick up all double crossovers, and these are then referred to the proper single crossover classes, the "piece-by-piece" per cent. estimate, and the percentage obtained from the cross, are found in complete agreement.

The amount of double crossing over in *Drosophila* is so large that the percentage of "crossing over" is rarely or

never more than 50 per cent., although the actual numbers given for "distances" between two genes may be as much as 107 based on summation of short distances. The latter method of calculation is the accurate way of stating the result, and whenever possible it is adhered to, *i.e.*, the percentage numbers for crossing over are sum totals based on results obtained with genes so near together that double crossing over is practically excluded.

Another illustration where the difference between the direct calculation between two factors (scute and forked) and the "piece-by-piece" estimate is greater than 50, is as follows: At one end of the series of sex-linked genes is a factor scute (zero) and near the other end forked. The direct data for crossing over between them gave a crossover value of 48.2. Between them three other loci were present in the same experiment, and crossing over between them could also be detected. As shown in the table below, the sum of these crossover values gave 61.1 units between scute and forked.
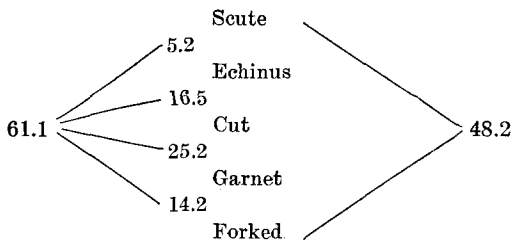
The presence of the intermediate factors makes it possible to pick up most of the double crossing over that occurred between scute and forked. When a correction is made for these the difference between 48.2 and 61.1 entirely disappears. Another and still more extreme example will help to make this more evident. Near one end of the second chromosome is the gene for star (eyes),

near the other end is the gene for speck. Bridges furnishes the following data in regard to crossing over between these loci. When only these distant loci are used the crossover value is 48.7. When the sum of the *crossover values between the following seven genes is* taken as the value for star and speck it amounts to 104.4.
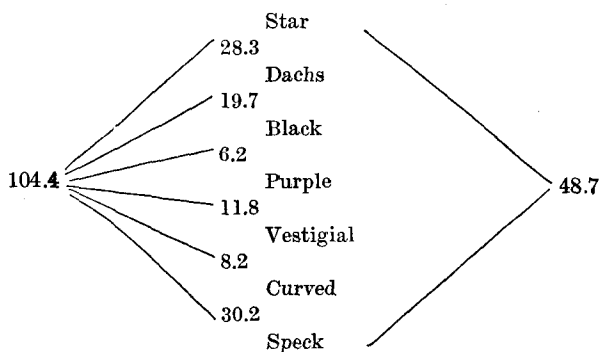

In this case (as other experiments show) there are still two units missing in the map distance given above (104.4), because of one per cent. of double crossovers in the region between curved and speck, that are not here recorded, since there are here no loci within this distance of 30.2.

Whenever cases in which double crossing over has taken place are checked up as in the foregoing cases, it is found that the discrepancies in the two methods are accounted for.

It is instructive to compare the preceding case with another one including several of the same genes, but in addition something else (deficiency), that cuts down the amount of crossing over in certain regions. The crossover value between star and speck was found to be 47.7 in this experiment. The sum of the crossover values of the six loci involved gave 79.3 units. The difference

between 47.7 and 79.3 is due to double crossovers as the data for the intermediate regions show:
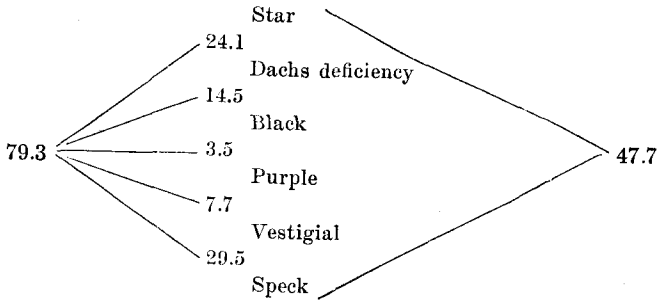
There was known to be present in this case a factor called "deficiency" in one of the two second chromosomes involved. It is near "dachs" and cuts down the crossing over between star and speck by about 25.8 units. It will be noticed that while the second summation value for star speck, *viz.*, 79.3, and the first value, *viz.*, 104.4, are very different, the crossover value between star and speck is in one case 47.7 and in the other 48.7. The meaning of this, as shown by data for intermediate loci, is that by the addition of 25.1 units (104.4 minus 79.3 equals 25.1) the number of double crossovers has so greatly increased that a difference of only about 1 per cent. of apparent crossing over is recorded in the star speck value.

"Distance" and Linear Order

The linear order of the genes implies distance between them, for which the crossover values stand as indices. It is obvious that if the order of the genes remained the same but something doubled the number of crossovers between two loci, their "distance" apart would at the same time appear to have been doubled. Again, if crossing over is thought of as due to twisting of the chromosomes of a pair about each other, then if the twisting is more likely to occur at the ends of the chromosomes, or if the twists themselves

are shorter there, "distance" in these regions is on a different scale from distance in the middle of the same chromosome.   Factors for crossing over have been found, by Sturtevant, that change the values in certain parts of the series and leave other parts unaffected.   When the influence of these special genes, that can be treated in the same way as are all Mendelian genes, is removed, the region that was affected gives its original crossover values again.

It is to be understood, then, that when we substitute the idea of distance for crossing over values the term is not used in an absolute sense, but in a relative sense, and that it depends always on the conditions of the experiment. That the genes do stand at definite levels in the chromosomes, and that in this sense they are definitely spaced, seems reasonable in the light of all the evidence bearing on this point; but even if they are so spaced that crossing over is a function of their distance from each other in the series, any influence that determines how often interchange between homologous pairs will take place would give the appearance that the actual distances themselves have changed.