The mechanics of crossing-over

BY K. V. Srinath, Botany Department, Central College, Bangalore

(Communicated by H. J. Bhabha, F.R.S.—Received 21 September 1948)

A theoretical treatment of the frequency distribution of chiasmata along a chromosome is developed, based on Darlington's postulate of relating crossing-over to chromosome coiling. Each chromatid is assumed to be an elastic thread subject to torsion and longitudinal stress, which is liable to break (but not yield) at a definite breaking load. It is supposed further that when an interchange occurs after a break, the consequent relief of torsional strain is confined to a small region round the location of the break, resulting in a lowering of the probability of a further break. The assumption about the effect of a break is formulated in a suitable way and an expression for the distribution function is thence derived. It is shown that in the limit of no interference this distribution goes over to the Poisson type, in conformity with Haldane's conclusions. Expressions are calculated for the mean and the standard deviation of the distribution; the former gives the relation between the mean chiasma frequency and chromosome-length.

When homologous chromosomes pair in meiosis, there is a mutual shift or interchange of segments. This 'crossing-over' is a biological event of fundamental importance, and many theories have been offered which attempt to explain the precise mechanism involved during this process. According to an attractive theory proposed by Darlington (1935) crossing-over is related to chromosome coiling. Two homologous chromosomes which are unwinding their relic coils pair in such a way that their internal and relational coils reach equilibrium. As a result of division of each chromosome into two thinner chromatids the strain of coiling is supposed to lead to a break in one of the chromatids. The consequent readjustment of forces results in a non-identical chromatid breaking at the same level. The broken ends of one chromatid are then very likely to meet those of the other and this produces crossing-over. The local reduction of coiling stresses involves a lowering of the possibility of a further break in that region, resulting in interference.

The theory of Darlington outlined above is satisfactory as affording a qualitative explanation of the various aspects of the event. But a quantitative approach will require a more precise specification of the nature of a chromosome in relation to its torsional and breaking properties. To this end it is supposed that each chromatid may be treated as an elastic thread subject to torsion and longitudinal stress (load). The assumption is made that for each thread there is a definite load (breaking load) at which it breaks, but does not yield. In an earlier contribution (Srinath 1946) it was shown how, on the basis of the elastic theory, one may derive a proof that when one of the chromatids breaks the other also experiences a break at the same level. This paper embodies a further elaboration of the elastic theory, and on the basis of suitable assumptions, a distribution function is derived for multiple crossing-over, taking explicit account of the effect of interference. It will appear that in the
absence of interference, the distribution reduces to the Poisson distribution, and this is in agreement with Haldane's conclusions (1931).

As indicated above, each chromatid is supposed to be an elastic thread with torsion, and that the two threads are in equilibrium under the torsional and longitudinal stresses. When an interchange occurs after a break, it is supposed that the consequent relief of torsional strain is confined to a small length \( \epsilon/2 \) on either side of the break. It is postulated that when a break has occurred, a further break cannot occur within a length \( \epsilon/2 \) on either side of the location of the break. Further, it is also assumed that \( \epsilon \) is small in comparison with total length \( l \) of the chromosome. The occurrence of each break results in a definite lowering of the probability of the next break.

Since the two chromatids involved in an interchange may be supposed to have identical elastic properties, we may, for the purpose of mathematical formulation, treat the problem as that of a single elastic thread along which breaks occur. After each break the ends are supposed to reunite again, resulting in the exclusion of a further break, in a region of length \( \epsilon \) round the location of the break and in a lowering of the probability for the next break.

The assumption about the effect of the occurrence of a break on the probability of occurrence of the next break may be formulated more precisely as follows: If \( \sigma \) is the breaking load and \( f_r(\sigma) \) the probability-density of \( \sigma \) after the \( r \)th break, then

\[
\int_0^S f_r(\sigma) \, d\sigma
\]

is the chance of the thread not exceeding \( S \) in strength, after the \( r \)th break. Evidently this chance diminishes with increasing \( r \), in conformity with the assumption made here that each break lessens the probability of the next. Specifically, it is assumed that

\[
\int_0^S f_{r+1}(\sigma) \, d\sigma = [(l-rc)/l] \int_0^S f_r(\sigma) \, d\sigma. \tag{1}
\]

According to the assumptions made, each break eliminates a segment \( \epsilon \) of the length available for further breaks. These are equivalent to assuming that the chance of the thread not exceeding \( S \) in strength is reduced at each break in the ratio of the available effective length of the thread to the total length. The reduction-factor \((l-rc)/l\) in (1) may also be rendered plausible by the following argument: if we assume this factor to be of the form \( g(\epsilon) \), we may, on account of the smallness of \( \epsilon \), write

\[
g(\epsilon) = a + b\epsilon.
\]

In the absence of interference, i.e. for \( \epsilon = 0 \), we must have \( g(0) = 1 \). Thus \( a = 1 \) and, since \( g \) must be non-dimensional, we must put \( b = -k/l \). This gives an expression of the above form. Since (1) is true for any \( S \), it is equivalent to

\[
f_{r+1}(\sigma) = \frac{l-rc}{l} f_r(\sigma). \tag{2}
\]

We are now in a position to determine the probability of \( n \) breaks. If \( \sigma_1, \sigma_2, \ldots, \sigma_n \) are the strengths of the thread at the 1st, 2nd, \ldots, \( n \)th breaks, under the longitudinal
stress $S$, then the probability $\phi_n$ of $n$ breaks is following a procedure due to Daniels (1945):

$$
\phi_n = P \int_0^S f_1(\sigma_1) d\sigma_1 \int_{\sigma_1}^S f_2(\sigma_2) d\sigma_2 \ldots \int_{\sigma_{n-1}}^S f_n(\sigma_n) d\sigma_n \\
= P \frac{(l-e)(l-2e)\ldots(l-ne)}{L^n} \int_0^S f_1(\sigma_1) d\sigma_1 \int_{\sigma_1}^S f_2(\sigma_2) d\sigma_2 \ldots \int_{\sigma_{n-1}}^S f_1(\sigma_n) d\sigma_n \\
= \frac{(l-e)(l-2e)\ldots(l-ne)}{L^n} P \frac{F(S) - F(0)}{L^n},
$$

where $P$ is a factor to be determined presently and $F(x) = \int^x f_1(x) dx$.

If we set $\lambda = F(S) - F(0)$,

we get

$$
\phi_n = \frac{(l-e)(l-2e)\ldots(l-ne)}{L^n} P \frac{\lambda^n}{L^n}. \quad (3)
$$

To determine $P$ the following condition is used:

$$
\sum_0^\infty \phi_n = 1.
$$

This gives

$$
P = \left(1 + \frac{\lambda e}{l}\right)^{-\left[\frac{l}{\epsilon}\right]+1}
$$

so that, finally,

$$
\phi_n = \frac{(l-e)(l-2e)\ldots(l-ne)}{L^n} \left(1 + \frac{\lambda e}{l}\right)^{-\left[\frac{l}{\epsilon}\right]+1} \frac{\lambda^n}{L^n}. \quad (4)
$$

Actually $n$ can only take the values $0, 1, 2, \ldots, N$, where $N = [l/e]$ is an integer such that $N \leq l/e < N + 1$. The probability $\phi_n$ is then conditioned by $\sum_0^N \phi_n = 1$. Thus

$$
P \sum_0^N \frac{(l-e)(l-2e)\ldots(l-ne) \lambda^n}{L^n} = 1,
$$

i.e.

$$
P \sum_0^N \frac{\left(\frac{l}{e}-1\right)\left(\frac{l}{e}-2\right)\ldots\left(\frac{l}{e}-n\right) \left(\frac{\epsilon \lambda}{l}\right)^n}{L^n} = 1.
$$

Let $l/e = N + \alpha$, $0 \leq \alpha < 1$. If $\epsilon \ll l$, $N$ will be large, and $\alpha$ small. Then the above condition becomes

$$
P \sum_0^N \frac{(N+\alpha-1)(N+\alpha-2)\ldots(N+\alpha-n) \left(\frac{\epsilon \lambda}{l}\right)^n}{L^n} = 1.
$$

Since $N$ is large and $\alpha$ small, we may neglect $\alpha$ in each factor in the numerator. This gives

$$
P \sum_0^N \frac{(N-1)(N-2)\ldots(N-n) \left(\frac{\epsilon \lambda}{l}\right)^n}{L^n} = 1,
$$

or

$$
P \sum_0^{N-1} \frac{(N-1)(N-2)\ldots(N-n) \left(\frac{\epsilon \lambda}{l}\right)^n}{L^n} = 1.
$$
Thus, finally, we have
\[ P \left( 1 + \frac{e\lambda}{l} \right)^{N-1} = 1 \quad \text{or} \quad P = (1 + e\lambda/l)^{-N+1} \]

and the expression for the probability \( \phi_n \) becomes
\[ \phi_n = \frac{(l-e)(l-2e)\ldots(l-ne)}{ln} \left( 1 + \frac{\lambda e}{l} \right)^{-U(e/l+1)} \frac{\lambda^n}{L^n} \]
(5)

For \( e \ll l \), this practically coincides with (4). The corresponding generating function \( \psi(t) \) is given by
\[ \psi(t) = \sum_{n=0}^{N} \phi_n t^n = \left( 1 + \frac{e\lambda}{l} \right)^{-N+1} \left( 1 + \frac{e\lambda}{l} \right)^{N-1} \]
(6)

From (5) we see that, as \( e \to 0 \),
\[ \phi_n \to e^{-\lambda} \frac{\lambda^n}{L^n} \]
(7)

which is the Poisson distribution, so that in the absence of interference, the distribution reduces to the Poisson type. It will also be noticed that the same conclusion would follow from (5) if, keeping \( e \) fixed, we let \( l \to \infty \). In other words there should be no interference in an infinitely long thread. This is in conformity with our general expectations, since, in an infinitely long thread, a break can always occur at a sufficient distance from previous ones, so as to avoid the effect of interference of these. The fact that independence from interference is attained for sufficiently long chromosomes is a well established experimental result Muller (1916).

We proceed to calculate the mean and standard deviation for the distribution (5). The mean \( m \) is given by
\[ m = \sum_{n=0}^{N} n\phi_n = \left[ \frac{d}{dt} \psi(t) \right]_{t=1} \]
\[ = \left( 1 + \frac{e\lambda}{l} \right)^{-N+1} (N-1) \left( 1 + \frac{e\lambda}{l} \right)^{N-2} \frac{e\lambda}{l} \]
\[ = (N-1) \frac{(e\lambda/l)}{1 + (e\lambda/l)} . \]

Writing \( N = l/e \), approximately, we have
\[ m = \frac{\lambda \left( 1 - \frac{e}{l} \right)}{1 + \frac{e\lambda}{l}} . \]
(8)

Similarly the standard deviation \( s^2 \) is given by
\[ s^2 = \sum_{n=0}^{N} (m-n)^2 \phi_n \]
\[ = m^2 \sum_{n=0}^{N} \phi_n - 2m \sum_{n=0}^{N} n \phi_n + \sum_{n=0}^{N} n^2 \phi_n \]
\[ = m^2 + \sum_{n=0}^{N} n^2 \phi_n . \]
Now
\[ \sum n^z \phi_n = \left[ \frac{d}{dt} \left( t \frac{d\psi}{dt} \right) \right]_{t=1} \]
\[ = \frac{\epsilon \lambda}{l} \left( \frac{l-1}{\epsilon} \right) \left[ 1 + \lambda \left( 1 - \frac{\epsilon}{l} \right) \right] \] with \( N \sim \frac{l}{\epsilon} \).

and hence
\[ s^2 = \frac{\lambda \left( 1 - \frac{\epsilon}{l} \right)}{\left( 1 + \frac{\epsilon \lambda}{l} \right)^2}. \]  \hspace{1cm} (9)

For \( \epsilon \to 0 \), \( m = s^2 = \lambda \), so that the variance and the mean coincide, which is a characteristic of the Poisson distribution. The equations (8) and (9) put in evidence the effect of interference on the mean and the variance, which are connected by the relation
\[ m^2 = \lambda \left( 1 - \frac{\epsilon}{l} \right) s^2. \]  \hspace{1cm} (10)

The relation (8) also gives the chromosome length mean chiasma frequency relationship. The theoretically derived curve (Mather 1938), however, shows a departure from the experimental curve, the reason for which is under examination. This contribution is offered as a first attempt at a theoretical formulation of a difficult biological problem. Grateful thanks are due to Mr U. R. Thiruvengkatadass for his valuable assistance in working out this problem.

References

Haldane, J. B. S. 1931 Cytologia, Tokyo, 3, 54–65.