A MATHEMATICAL THEORY OF CHROMOSOMAL REARRANGEMENTS*

BY J. B. S. HALDANE, F.R.S., *University College, London*
AND D. E. LEA, *The Strangeways Laboratory, Cambridge*

(With Two Text-figures)

When a nucleus is exposed to X-rays, gene mutations and minute inversions and deficiencies are produced roughly in proportion to the dose. Large rearrangements, including inversions and translocations, occur with a frequency increasing more rapidly than the first power of the dose but (in *Drosophila*) less rapidly than its square (Muller, 1940, 1941). In the case of *Drosophila* spermatozoa a given dose in roentgens gives the same result regardless of its intensity or of intervals during its application. In other cases this is not so, since fragments appear to join during irradiation. It is generally agreed that the number of breaks primarily produced in the chromosomes of a nucleus is proportional to the dose (e.g. Sax, 1940; Carlson, 1941), and that large rearrangements are due to unions of two or more breaks. In *Drosophila*, at any rate, broken ends rarely unite with normal ends, and terminal deficiencies are also rare, though they seem to occur more frequently in *Zea* (Stadler, 1941).

The theory developed in this paper rests on the following assumptions (cp. Pontecorvo, 1942; Lea & Catcheside, 1945). We suppose that after a dose D of radiation, an average of αD breaks per nucleus are produced, made up of an average number αpD breaks which undergo sister-union and behave as dominant lethals in the manner envisaged by Pontecorvo and Muller, and an average number αqD breaks which do not undergo sister-union. $p+q=1$, and α, p and q are assumed to be constants which are independent of the dose (though possibly dependent on the state of the nucleus at the time of irradiation); and union among the breakage ends which do not undergo sister-union is presumed to be random. The assumption of random union of breakage ends will not be a valid approximation in all cases. It is fairly certain, for example, that in *Tradescantia* microspores union of breakage ends is far from random, but is mainly limited to breakage ends separated at the time of breakage by a distance less than one-tenth of the nuclear diameter (Lea & Catcheside, 1942). There is reason for believing, however, that union may be more nearly random in some other cases, particularly after irradiation of *Drosophila* sperm. A preference for restitutional union as opposed to unions resulting in chromosome structural changes is not incompatible with the present calculations, providing it can be assumed that the mean number of such preferentially restituting breaks is proportional to the dose. In this event αqD is interpreted to be the number of randomly uniting breaks remaining after subtraction of the preferentially restituting class of break.

* This paper was first drafted (by J.B.S.H.) in 1942, but was not published as he was unable to evaluate the factor q. Subsequently Lea & Catcheside (1945) working on the same problem obtained results which were, in part, identical.

Their calculations, however, involved the assumption that every break is in a different chromosome arm, which is strictly only a valid approximation when the number of arms is large compared with the number of breaks in the nucleus. It has been thought worth while, therefore, to complete the present paper, since the method it describes is not subject to this limitation, but can be applied to nuclei with any given number of arms, as well as to nuclei with many arms.

Rearrangements which are *eucentric* (or *symmetrical*), that is to say in which all the chromatin is attached to a centromere and in which therefore none is attached to two centromeres, will be viable. Arrangements which are *dyscentric* (or *asymmetrical*) will make a further contribution to the dominant lethals.

Rearrangement after r breaks

In this section we limit attention to nuclei in which no breaks undergo sister-union. First consider the total number of possible arrangements after r breaks. There are $2r$ breakage ends to unite in pairs. The first can unite with any of $2r-1$ partners, the third with any of $2r-3$, and so on. Thus the number of possible arrangements is

$$1.3.5.....(2r-3)\,(2r-1)=\frac{(2r)!}{2^r r!}. \tag{1}$$

This is so regardless of whether the breaks occur in the same or in different chromosomes. We have assumed that all arrangements are equally probable.

But only some are eucentric, including of course the original arrangement. The number of eucentric arrangements varies with the structure of the nucleus. We shall show that it is minimal when all the breakage ends are in different chromosomes or different arms of metacentric or V-shaped chromosomes, and maximal when all are in one (or two) telocentric or rod-shaped chromosomes, or in one (or two) arms of V-shaped chromosomes.

If the r breaks are all in different arms, there are r acentric breakage ends to be united with r breakage ends connected with the centromeres. Clearly they can be attached in $r!$ ways. Thus, if the number of chromosomes is large compared with r, a condition often fulfilled in practice, the frequency of eucentric arrangements is

$$\frac{1.2.3.....(r-1)\,r}{1.3.5.....(2r-3)\,(2r-1)}=\frac{2^r(r!)^2}{(2r)!}.$$

Since one of the $r!$ eucentric arrangements restitutes the original nucleus, the frequency of rearrangements giving eucentric nuclei with structural change is $\dfrac{2^r r!\,(r!-1)}{(2r)!}$.

If the r breaks all occur in one arm, there are $(r-1)$ fragments to be filled in between the centromere and the free end. They can be filled in in $(r-1)!$ orders. And each can be in the original direction, or inverted. Thus the total number of arrangements is $2^{r-1}(r-1)!$. This is equal to $r!$ if $r=1$ or 2, and exceeds it if $r>2$.

Now consider r breaks distributed over k arms. k terminal fragments are separated from the centromeres, and $r-k$ interstitial fragments are separated from centromeres and ends. If the centromeres and segments containing them are regarded as fixed we can arrange the k chromosome ends in relation to them in $k!$ ways, leaving k gaps. Into these k gaps, or some of them, $r-k$ fragments each with two breakage ends are to be fitted. The first can be put into any of the k gaps leaving $k+1$ gaps, the second into any of the $k+1$ gaps leaving $k+2$ gaps, and so on. Further, each of these $r-k$ segments can be placed in either of two directions. Thus the total number of arrangements is

$$k!\,k(k+1)\,(k+2)...(r-1)\,2^{r-k}=2^{r-k}(r-1)!\,k. \tag{2}$$

This expression has been given by Fano (1942). It is equal to $2^{r-1}(r-1)!$ if $k=1$ or 2, otherwise it lies between $2^{r-1}(r-1)!$ and $r!$. To sum up: if two breaks have occurred, the probability of a eucentric arrangement is $\frac{2}{3}$. If r breaks have occurred, the probability is

$\frac{2^{2r-1}\,r!\,(r-1)!}{(2r)!}$ for a nucleus containing one or two arms, and approximates to $\frac{2^r(r!)^2}{(2r)!}$ for one containing a large number of arms. For nuclei with an intermediate number of arms, the probability is intermediate.

We have now to calculate this intermediate probability, and also the frequency of inversions in a given chromosome. The former can be done with ease for the case where all the arms are equal. Consider the chromosome set consisting of three arms which are equal in the sense that the probability of a break in each is equal. If there are three breaks, the probability that they will all be in different arms ($k=3$) is $\frac{2}{9}$. In this case, the number of eucentric arrangements is 3! or 6. The probability that all the breaks are in one or two arms is $\frac{7}{9}$. In this case the number of eucentric arrangements is $2^2 2!$ or 8. Thus the mean number is $6\times\frac{2}{9}+8\times\frac{7}{9}=7\frac{5}{9}$, and since the total number of arrangements is 15, the frequency of eucentric arrangements is $\frac{68}{135}$. The mean number of eucentric rearrangements, i.e. of arrangements differing from the original, is $6\frac{5}{9}$, and their frequency is $\frac{59}{135}$.

In general, if $p(a, r, k)$ is the probability that all of r breaks will occur in k out of a total of a equal arms, $p(a, r, k)$ is a^{-r} times the number of terms in the expansion of

$$(x_1+x_2+\ldots+x_a)^r$$

which contain just k different x's. The expansion can be made by the multinomial theorem; alternatively tables are available facilitating the writing down of $p(a, r, k)$ for any set of values of a, r and k (Fisher & Yates, 1938, Table XXII, Initial differences of powers of natural numbers). For example, if there are five equal arms and six breaks, $a=5$, $r=6$, and we find $p(5,6,1)=\frac{1}{3125}$; $p(5,6,2)=\frac{124}{3125}$; $p(5,6,3)=\frac{216}{625}$; $p(5,6,4)=\frac{312}{625}$; $p(5,6,5)=\frac{72}{625}$.

The mean number of eucentric arrangements is

$$E_r=(r-1)!\sum_{k=1}^{a} p(a, r, k)\,2^{r-k}k.$$

Thus for five equal arms and six breaks the mean number of eucentric arrangements is

$$\frac{5!}{625}(25.2^4.2+216.2^3.3+312.2^2.4+72.2.5)=2245{\cdot}632.$$

The mean number of eucentric rearrangements (i.e. arrangements excluding the original one) is one less, or 2244·632.

In general, if there are r breaks in a chromosome set containing a equal arms, we proceed as follows. Of the a^r terms in the expansion of $(x_1+x_2+\ldots+x_{a-1}+x_a)^r$, a contain just one letter. Thus $p(a, r, 1)=a^{-r}.a$. The number of terms containing x_1 or x_2 or both, but none of $x_3, x_4, \ldots, x_a$ is obtained by putting $x_1=x_2=1$, $x_3=x_4=\ldots=x_a=0$, and is therefore 2^r. Subtracting the two terms (x_1^r and x_2^r) which contain x_1 or x_2 separately, we find that there are 2^r-2 terms which contain both x_1 and x_2 but no other letter. Hence there are $\binom{a}{2}(2^r-2)$ terms which contain just two letters, $\binom{a}{2}=\frac{1}{2}a\,(a-1)$ being the number of ways of selecting two letters from the total of a letters. Thus $p(a, r, 2)=a^{-r}\binom{a}{2}(2^r-2)$.

Again, there are 3^r terms which contain any or all of x_1, x_2, x_3, but none of $x_4, x_5, \ldots, x_a$. Subtracting 3 for the terms in x_1, x_2 or x_3 alone, and $3(2^r-2)$ for the terms in x_1 and x_2, x_2 and x_3, x_3 and x_1, we find that there are $3^r-3.2^r+3$ terms which contain all of x_1, x_2, x_3,

and none of $x_4, x_5, \ldots, x_a$. Thus $p(a, r, 3) = a^{-r}\binom{a}{3}(3^r - 3.2^r + 3)$, $\binom{a}{3} = \frac{1}{6}a(a-1)(a-2)$ being the number of ways of selecting 3 letters from the total of a. Thus

$$E_r = \frac{(r-1)!}{a^r}\left\{\binom{a}{1}2^{r-1} + \binom{a}{2}(2^r-2)2^{r-2}2 + \binom{a}{3}(3^r-3.2^r+3)2^{r-3}3 + \binom{a}{4}(4^r-4.3^r+6.2^r-4)2^{r-4}4 + \ldots\right\} \quad (3)$$

where the summation extends over a terms (terms after the rth vanish if $r < a$). In this way we find:

$$\left.\begin{aligned}
&\text{For 1 or 2 arms:} && E_r^{\mathrm{I}} = E_r^{\mathrm{II}} = 2^{r-1}(r-1)!.\\
&\text{For 3 equal arms:} && E_r^{\mathrm{III}} = 3.2^{r-3}(r-1)!\left\{1 + \frac{2^r-1}{3^r}\right\}.\\
&\text{For 4 equal arms:} && E_r^{\mathrm{IV}} = 4.2^{r-4}(r-1)!\left\{1 + \frac{2.3^r-2}{4^r}\right\}.\\
&\text{For 5 equal arms:} && E_r^{\mathrm{V}} = 5.2^{r-5}(r-1)!\left\{1 + \frac{3.4^r+2.3^r-2.2^r-3}{5^r}\right\}.\\
&\text{For many arms:} && E_r^{\infty} = r!.
\end{aligned}\right\} \quad (4)$$

The superscript attached to E_r indicates the number of arms in the chromosome set. The values of E_r for these different types of chromosome set have been tabulated in Table 1 up to $r = 24$. Calculations could be made for unequal arms if the probability that a break should occur in a given arm were known. In many *Drosophila* species there are four nearly equal arms, three or four being autosomal, a somewhat longer arm constituting the whole or part of the X, and usually a very small autosomal arm. Thus the values of E_r for an X-bearing spermatozoon will be fairly close to those calculated for five equal arms.

Also tabulated in Table 1 is $\frac{(2r)!}{r!\,2^r}$ which is the total number of arrangements, eucentric and dyscentric, of r breaks. For large values of r the ratio of E_r to the total number of arrangements is approximately $\frac{a}{2^a}\sqrt{\binom{\pi}{r}}$ (Fano, 1942).

We can also calculate the probability of an eucentric arrangement within an arm, e.g. within the X-chromosome of *Drosophila melanogaster*. Taking the probability of a break in the X-chromosome to be one-fifth of the probability of a break anywhere in the set, the probability that there will be l breaks in the X-chromosome out of a total of r breaks in the set is $\binom{r}{l}4^{r-l}5^{-r}$, where $\binom{r}{l} = \frac{r!}{l!\,(r-l)!}$. These l breaks can unite eucentrically in E_l^{I} ways without any part of the X-chromosome being translocated or any part of any other chromosome being united to it. One way is complete restitution, $(E_l^{\mathrm{I}} - 1)$ ways are inversions. The other $r-l$ breaks occur among the other four equal arms. They can unite eucentrically in E_{r-l}^{IV} ways (where E_{r-l}^{IV} may be read for any value of $(r-l)$ in Table 1). Hence the mean number of eucentric arrangements in which the X-chromosome suffers no structural change is

$$H_r^{\mathrm{V}} = 5^{-r}\sum_{l=0}^{r}\binom{r}{l}4^{r-l}E_{r-l}^{\mathrm{IV}}, \quad (5)$$

Table 1

r	$\frac{(2r)!}{r!\,2^r}$	$E_r^{\mathrm{I}}=E_r^{\mathrm{II}}$	E_r^{III}	E_r^{IV}	E_r^{V}	E_r^{∞}	H_r^{V}	I_r^{V}
0	1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1	0
2	3	2	2	2	2	2	1·64	0·04
3	1·5 $\times 10^1$	8	7·5556	7·25	7·04	6	4·584	0·152
4	1·05 $\times 10^2$	4·8 $\times 10^1$	4·2667 $\times 10^1$	3·9 $\times 10^1$	3·648 $\times 10^1$	2·4 $\times 10^1$	1·9278 $\times 10^1$	0·5616
5	9·45 $\times 10^2$	3·84 $\times 10^2$	3·2474 $\times 10^2$	2·8275 $\times 10^2$	2·5405 $\times 10^2$	1·2 $\times 10^2$	1·1022 $\times 10^2$	2·6250
6	1·0395 $\times 10^4$	3·84 $\times 10^3$	3·1289 $\times 10^3$	2·6025 $\times 10^3$	2·2456 $\times 10^3$	7·2 $\times 10^2$	8·0362 $\times 10^2$	1·6020 $\times 10^1$
7	1·3514 $\times 10^5$	4·608 $\times 10^4$	3·6567 $\times 10^4$	2·9188 $\times 10^4$	2·4218 $\times 10^4$	5·04 $\times 10^3$	7·1588 $\times 10^3$	1·2417 $\times 10^2$
8	2·0270 $\times 10^6$	6·4512 $\times 10^5$	5·0264 $\times 10^5$	3·8714 $\times 10^5$	3·0958 $\times 10^5$	4·032 $\times 10^4$	7·5552 $\times 10^4$	1·1785 $\times 10^3$
9	3·4459 $\times 10^7$	1·0322 $\times 10^7$	7·9424 $\times 10^6$	5·9359 $\times 10^6$	4·5877 $\times 10^6$	3·6288 $\times 10^5$	9·2291 $\times 10^5$	1·3269 $\times 10^4$
10	6·5473 $\times 10^8$	1·8579 $\times 10^8$	1·4176 $\times 10^8$	1·0336 $\times 10^8$	7·7453 $\times 10^7$	3·6288 $\times 10^6$	1·2815 $\times 10^7$	1·7292 $\times 10^5$
11	1·3749 $\times 10^{10}$	3·7159 $\times 10^9$	2·8191 $\times 10^9$	2·0149 $\times 10^9$	1·4688 $\times 10^9$	3·9917 $\times 10^7$	1·9934 $\times 10^8$	2·5587 $\times 10^6$
12	3·1623 $\times 10^{11}$	8·1750 $\times 10^{10}$	6·1785 $\times 10^{10}$	4·3464 $\times 10^{10}$	3·0924 $\times 10^{10}$	4·7900 $\times 10^8$	3·4330 $\times 10^9$	4·2345 $\times 10^7$
13	7·9059 $\times 10^{12}$	1·9620 $\times 10^{12}$	1·4791 $\times 10^{12}$	1·0276 $\times 10^{12}$	7·1584 $\times 10^{11}$	6·2270 $\times 10^9$		
14	2·1346 $\times 10^{14}$	5·1012 $\times 10^{13}$	3·8390 $\times 10^{13}$	2·6415 $\times 10^{13}$	1·8069 $\times 10^{13}$	8·7178 $\times 10^{10}$		
15	6·1903 $\times 10^{15}$	1·4283 $\times 10^{15}$	1·0737 $\times 10^{15}$	7·3325 $\times 10^{14}$	4·9389 $\times 10^{14}$	1·3077 $\times 10^{12}$		
16	1·9190 $\times 10^{17}$	4·2850 $\times 10^{16}$	3·2186 $\times 10^{16}$	2·1854 $\times 10^{16}$	1·4529 $\times 10^{16}$	2·0923 $\times 10^{13}$		
17	6·3327 $\times 10^{18}$	1·3712 $\times 10^{18}$	1·0294 $\times 10^{18}$	6·9591 $\times 10^{17}$	4·5759 $\times 10^{17}$	3·5569 $\times 10^{14}$		
18	2·2164 $\times 10^{20}$	4·6621 $\times 10^{19}$	3·4989 $\times 10^{19}$	2·3573 $\times 10^{19}$	1·5369 $\times 10^{19}$	6·4024 $\times 10^{15}$		
19	8·2008 $\times 10^{21}$	1·6783 $\times 10^{21}$	1·2593 $\times 10^{21}$	8·4627 $\times 10^{20}$	5·4722 $\times 10^{20}$	1·2165 $\times 10^{17}$		
20	3·1983 $\times 10^{23}$	6·3777 $\times 10^{22}$	4·7847 $\times 10^{22}$	3·2091 $\times 10^{22}$	2·0621 $\times 10^{22}$	2·4329 $\times 10^{18}$		
21	1·3113 $\times 10^{25}$	2·5511 $\times 10^{24}$	1·9137 $\times 10^{24}$	1·2816 $\times 10^{24}$	8·1931 $\times 10^{23}$	5·1091 $\times 10^{19}$		
22	5·6386 $\times 10^{26}$	1·0715 $\times 10^{26}$	8·0370 $\times 10^{25}$	5·3764 $\times 10^{25}$	3·4225 $\times 10^{25}$	1·1240 $\times 10^{21}$		
23	2·5374 $\times 10^{28}$	4·7144 $\times 10^{27}$	3·5301 $\times 10^{27}$	2·3635 $\times 10^{27}$	1·4994 $\times 10^{27}$	2·5852 $\times 10^{22}$		
24	1·1926 $\times 10^{30}$	2·1686 $\times 10^{29}$	1·6266 $\times 10^{29}$	1·0865 $\times 10^{29}$	6·8730 $\times 10^{28}$	6·2045 $\times 10^{23}$		

Table 2

aqD	S_1	S_2					S_3		$1-S_1/S_2$					S_3/S_2	
No. of arms...		1 or 2	3	4	5	Many	5	Many	1 or 2	3	4	5	Many	5	Many
0·32	1·337	1·357	1·357	1·357	1·357	1·356	0·3966	0·3952	0·0146	0·0145	0·0144	0·0143	0·0140	0·2923	0·2914
0·50	1·543	1·596	1·595	1·594	1·594	1·592	0·7000	0·6942	0·0330	0·0325	0·0322	0·0320	0·0309	0·4391	0·4360
0·72	1·811	1·932	1·929	1·928	1·926	1·920	1·171	1·152	0·0627	0·0615	0·0606	0·0601	0·0572	0·6080	0·5996
0·98	2·151	2·405	2·398	2·393	2·389	2·373	1·907	1·849	0·1056	0·1030	0·1011	0·0998	0·0935	0·7979	0·7793
1·28	2·577	3·078	3·060	3·047	3·038	2·996	3·068	2·915	0·1626	0·1576	0·1541	0·1517	0·1396	1·010	0·9731
1·62	3·107	4·053	4·010	3·980	3·959	3·858	4·933	4·553	0·2334	0·2250	0·2191	0·2150	0·1945	1·246	1·180
2·00	3·762	5·502	5·403	5·334	5·287	5·050	7·986	7·090	0·3162	0·3037	0·2947	0·2885	0·2565	1·510	1·401
2·88	5·557	11·21	10·75	10·43	10·21	9·172	21·86	17·29	0·5041	0·4831	0·4671	0·4556	0·3941	2·142	1·885
3·92	8·263	26·57	24·59	23·15	22·17	17·78	65·25	43·23	0·6894	0·6643	0·6435	0·6278	0·5357	2·943	2·432
5·12	12·29	74·67	66·03	59·60	55·22	36·82	218·2	112·2	0·8355	0·8139	0·7938	0·7775	0·6663	3·951	3·046
6·48	18·31	250·9	211·7	181·4	160·8	81·57	834·8	304·6	0·9270	0·9135	0·8991	0·8861	0·7755	5·190	3·734
8·00	27·31	1009	817·6	662·6	557·3	193·6	3718	870·8	0·9729	0·9606	0·9588	0·9510	0·8590	6·671	4·497

while the mean number of eucentric rearrangements in which the X-chromosome suffers inversion but not interchange with autosomes is

$$I_r^{\mathrm{V}} = 5^{-r} \sum_{l=2}^{r} \binom{r}{l} 4^{r-l} E_{r-l}^{\mathrm{IV}} (E_l^{\mathrm{I}} - 1). \tag{6}$$

H_r^{V} and I_r^{V} are tabulated in Table 1 for values of r up to 12.

Rearrangement after a given dose

If αD is the mean number of breaks primarily produced per nucleus by a dose D, then $e^{-\alpha D}(\alpha D)^r/r!$ is the proportion of nuclei having r breaks, and $e^{-\alpha D}(\alpha q D)^r/r!$ is the proportion of nuclei having r breaks none of which undergoes sister-union. The following results then follow from the preceding section:

The proportion of nuclei which are eucentric and without aberrations is

$$X = e^{-\alpha D} S_1, \quad \text{where} \quad S_1 = 1 + \alpha q D + \tfrac{1}{2}(\alpha q D)^2 + \ldots + \frac{(2\alpha q D)^r}{(2r)!} + \ldots. \tag{7}$$

The proportion of nuclei which are eucentric, with or without aberration, is

$$Y = e^{-\alpha D} S_2, \quad \text{where} \quad S_2 = 1 + \alpha q D + \tfrac{1}{3}(\alpha q D)^2 + \ldots + \frac{(2\alpha q D)^r}{(2r)!} E_r + \ldots. \tag{8}$$

The total number of primary breaks formed in eucentric nuclei, per total nucleus is

$$Z = e^{-\alpha D} S_3, \quad \text{where} \quad S_3 = \alpha q D + \tfrac{2}{3}(\alpha q D)^2 + \ldots + \frac{(2\alpha q D)^r}{(2r)!} E_r r + \ldots. \tag{9}$$

The series S_1 is the same for different chromosome sets and has the sum

$$S_1 = \cosh \sqrt{(2\alpha q D)}, \tag{10}$$

where $\cosh x = \frac{1}{2}(e^x + e^{-x})$ is the hyperbolic cosine.

The series S_2 and S_3 are different for chromosome sets containing different numbers of arms, and where necessary we attach a superscript to indicate the number of arms, e.g. S_2^{V} means the value of S_2 when E_r in formula (8) takes the values appropriate to five equal arms.

An algebraic expression for the sums S_2 and S_3 may be obtained in the case of a chromosome set with many arms. Defining

$$y = \sum_{r=1}^{\infty} \frac{(2x)^{2r} r!}{(2r)!}, \tag{11}$$

and differentiating (y/x) term by term (which is permissible since the series is absolutely convergent for all values of x), we obtain

$$\frac{d}{dx}\left(\frac{y}{x}\right) - 2y = 2,$$

whence

$$\frac{d}{dx}(e^{-x^2} y/x) = 2e^{-x^2}.$$

Integrating,

$$e^{-x^2} y/x = 2\int_0^x e^{-x^2} dx = \sqrt{\pi}\ \mathrm{erf}\, x,$$

where erf x, defined as $\dfrac{2}{\sqrt{\pi}}\displaystyle\int_0^x e^{-x^2} dx$, is the error function.

Thus
$$y = x\sqrt{\pi}\, e^{x^2} \operatorname{erf} x. \tag{12}$$

Evidently $S_2^{\infty} = 1 + y$, if $x = \sqrt{(\frac{1}{2}\alpha q D)}$. Hence

$$S_2^{\infty} = 1 + \sqrt{(\tfrac{1}{2}\pi\alpha q D)}\, e^{\frac{1}{2}\alpha q D} \operatorname{erf} \sqrt{(\tfrac{1}{2}\alpha q D)}. \tag{13}$$

By differentiating (11) we obtain

$$\sum_{r=1}^{\infty} \frac{(2x)^{2r} r!\, r}{(2r)!} = \tfrac{1}{2}x \frac{dy}{dx}$$

and by differentiating (12) we obtain

$$\tfrac{1}{2}x \frac{dy}{dx} = x^2 \left\{ 1 + \frac{1+2x^2}{x} \frac{\sqrt{\pi}}{2} e^{x^2} \operatorname{erf} x \right\}.$$

Inserting $x = \sqrt{(\frac{1}{2}\alpha q D)}$ we obtain

$$S_3^{\infty} = \tfrac{1}{2}\alpha q D \left\{ 1 + \frac{1+\alpha q D}{\sqrt{(\frac{1}{2}\alpha q D)}} \frac{\sqrt{\pi}}{2} e^{\frac{1}{2}\alpha q D} \operatorname{erf} \sqrt{(\tfrac{1}{2}\alpha q D)} \right\}. \tag{14}$$

Equations (10), (13) and (14) have been given by Lea & Catcheside (1945). Adequate tables are available of the exponential, hyperbolic cosine, and error functions (e.g. *W.P.A. Tables*, 1939–41), which are used to evaluate S_1, S_2^{∞} and S_3^{∞} in Table 2.

The evaluation of S_2 in the case of chromosome sets with a small number of arms is less convenient. In the case of a chromosome set with one or two arms we may proceed as follows. From equations (11) and (12) we have

$$\frac{y}{x} = \sum_{r=1}^{\infty} \frac{2^{2r} x^{2r-1} r!}{(2r)!} = \sqrt{\pi}\, e^{x^2} \operatorname{erf} x.$$

Integrating between $x = 0$ and $x = \sqrt{(\alpha q D)}$ we have

$$\sum_{r=1}^{\infty} \frac{(2\alpha q D)^r\, 2^{r-1}\, (r-1)!}{(2r)!} = \sqrt{\pi} \int_0^{\sqrt{(\alpha q D)}} e^{x^2} \operatorname{erf} x\, dx,$$

i.e.
$$S_2^{\mathrm{I}} = 1 + \sqrt{\pi} \int_0^{\sqrt{(\alpha q D)}} e^{x^2} \operatorname{erf} x\, dx, \tag{15}$$

where, as before, $\operatorname{erf} x$ denotes $\dfrac{2}{\sqrt{\pi}} \displaystyle\int_0^x e^{-x^2}\, dx$.

With the aid of equations (4) it would be possible to derive somewhat more complicated expressions for the values of S_2^{III}, S_2^{IV} and S_2^{V} involving the integral which occurs in equation (15). However, tables of this integral are not available and it has been found more convenient to compute S_2^{I}, S_2^{III}, S_2^{IV}, S_2^{V} and S_3^{V} by direct summation of the series given in equations (8) and (9) than to perform the integration numerically. The first ten terms of the series suffice for values of $\alpha q D$ up to unity, twenty terms are needed to give five significant figures for a value of $\alpha q D$ as high as 6. The results of the computation are set out in Table 2.

COMPARISON WITH EXPERIMENT

Experiments on the irradiation of *Drosophila* gametes provide suitable means of comparing the predictions of the theory with experiment. One of the observable quantities is the proportion of viable sperm which have chromosome aberrations. The theoretical expression for this proportion is $1 - X/Y = 1 - S_1/S_2$, which is listed in Table 2 as a function of $\alpha q D$. Lea and Catcheside found that the theoretical curve for a chromosome set with

many arms fitted the experimental observations on *Drosophila melanogaster* satisfactorily when $\alpha q = 0{\cdot}57$ per 1000 r. This curve is shown in Fig. 1 as curve *A*, together with theoretical curves (*B* and *C*) calculated for a chromosome set with five equal arms, which should be more appropriate for *Drosophila*. Curve *B* is computed with the same value $\alpha q = 0{\cdot}57$ as was employed in curve *A*, to indicate the magnitude of the error involved in using the many-arm formula instead of the 5-arm formula. A better fit of the five-arm formula to the experimental points at 4000 and 5000 r. is obtained by taking a somewhat smaller value of αq, namely, $\alpha q = 0{\cdot}52$ per 1000 r., and this value has been used in computing curve *C*.

A second observable quantity is $1 - Y = 1 - e^{-\alpha D} S_2$, the proportion of total sperm which are non-viable. (In comparing this formula with experimental results on imago emergence with *Drosophila melanogaster*, the latter have to be corrected for the proportion of unirradiated eggs which fail to develop into adult flies.) Lea & Catcheside (1945) found that the many-arm formula satisfactorily agreed with experiment if the value $\alpha = 0{\cdot}75$ per 1000 r. was assumed, together with $\alpha q = 0{\cdot}57$ already derived. This curve is reproduced in Fig. 2 as curve *A*, together with the theoretical curve (*B*) calculated from the five-arm formula with the same constants $\alpha = 0{\cdot}75$ and $\alpha q = 0{\cdot}57$ per 1000 r. Again the difference between the two theoretical curves is not very great. A better agreement of the five-arm formula with the experimental points at large doses is obtained by combining the value $\alpha q = 0{\cdot}52$ used in curve *C* of Fig. 1 with $\alpha = 0{\cdot}78$ per 1000 r., and this has been done in computing curve *C* of Fig. 2. The values of α and q best fitting the experimental results on the two theories are set out in Table 3.

Table 3

	Many-arm formulae	Five-arm formulae
α = mean number of primary breaks per sperm per 1000 r.	0·75	0·78
αq = mean number of primary breaks per sperm per 1000 r. which do not undergo sister-union	0·57	0·52
p = probability of a break undergoing sister-union	0·24	0·33
q = probability of a break not undergoing sister-union	0·76	0·67

The changes made in the constants by the replacement of the many-arm formulae by the more appropriate five-arm formulae are not sufficient to upset seriously the agreement which Lea & Catcheside found between the value of q deduced in this manner and the value ($q = 0{\cdot}74$) deduced from an analysis of sex-ratio distortion (Catcheside & Lea, 1945*b*).

In connexion with their suggestion that recessive lethals accompany a certain proportion of chromosome breaks (including breaks which restitute) Lea & Catcheside (1945) pointed out that the number of recessive lethals (in viable sperm) should increase with dose in the same manner as S_3/S_2, since S_3/S_2 is the mean number of primary breaks per viable sperm. Calculating S_3 and S_2 on the many-arm formulae they found that S_3/S_2 increased approximately as the 0·84 power of the dose in the interval 1500–6000 r., which is not in accord with the experimentally found rather exact proportionality of the number of sex-linked recessive lethals induced by radiation to the dose. However, when S_3 and S_2 are computed on the five-arm formulae (see Table 2), it is found that S_3/S_2 increases approximately as the 0·92 power of the dose between 1500 and 6000 r., making the discrepancy less serious.

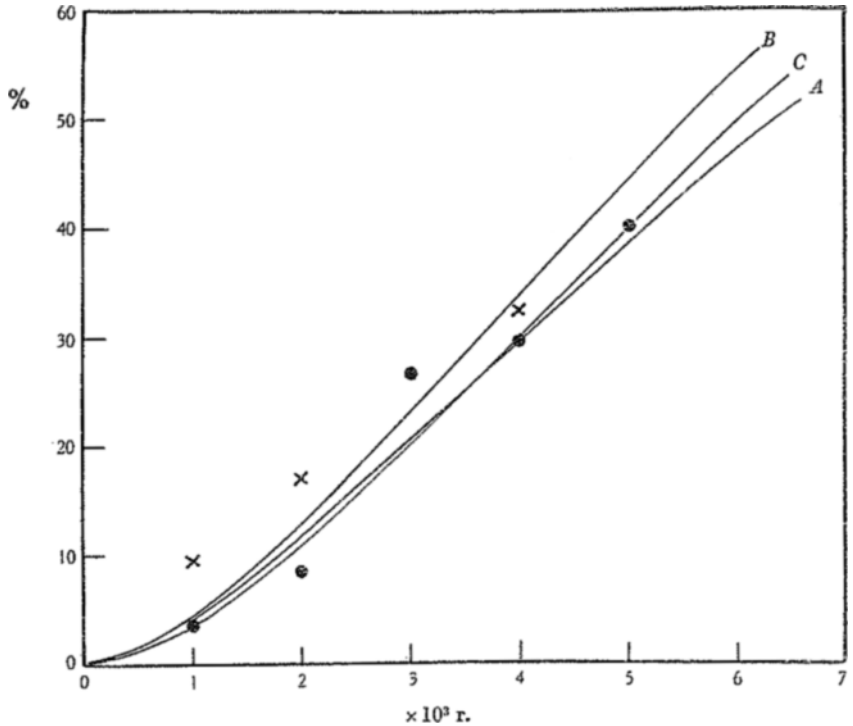

Fig. 1. Percentage of viable sperm having chromosomal structural change (curves theoretical, points experimental). ● Bauer, Demerec & Kaufmann (1938). × Catcheside (1938).

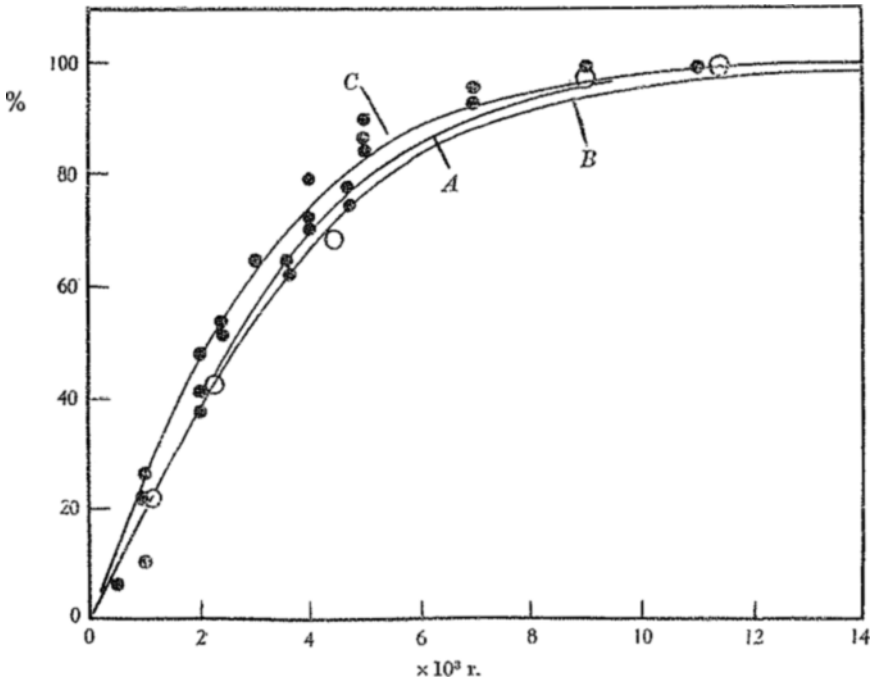

Fig. 2. Percentage of dominant lethals in *Drosophila melanogaster* sperm (curves theoretical, points experimental). ● Demérec & Fano (1944). ○ Catcheside & Lea (1945*a*).

The number of primary breaks per viable sperm (S_3/S_3) produced by a dose of 3000 r. Lea & Catcheside found to be 1·23, using the many-arm formulae with $\alpha q = 0{\cdot}57$ per 1000 r. When the calculation is repeated using the five-arm formulae and $\alpha q = 0{\cdot}52$ per 1000 r., the value is found to be 1·20. The agreement Lea & Catcheside found between estimates of the number of primary breaks per viable sperm based on the analysis of recessive lethals and of structural changes respectively is thus not upset by replacing their approximate calculations by the present more exact calculation.

Summary

A theory is given of the process of chromosomal structural rearrangement following irradiation. The theory applies to nuclei in which the union of breakage ends is at random, and assumes that the number of breaks primarily produced is proportional to the dose.

Formulae and tables are given enabling the proportion of nuclei undergoing eucentric and dyscentric types of rearrangement respectively to be calculated as a function of the dose, the cases of nuclei with 1, 2, 3, 4, 5 or many chromosome arms in the set being separately considered. The five-arm calculation is compared with published experimental results of the irradiation of the spermatozoa of *Drosophila melanogaster*.

REFERENCES

Bauer, H., Demerec, M. & Kaufmann, B. P. (1938). X-ray induced chromosomal alterations in *Drosophila melanogaster*. *Genetics*, **23**, 610.

Carlson, J. G. (1941). An analysis of X-ray induced single breaks in neuroblast chromosomes of the grasshopper (*Chortophaga viridifasciata*). *Proc. Nat. Acad. Sci., Wash.*, **27**, 42.

Catcheside, D. G. (1938). The effect of X-ray dosage upon the frequency of induced structural changes in the chromosomes of *Drosophila melanogaster*. *J. Genet.* **36**, 307.

Catcheside, D. G. & Lea, D. E. (1945*a*). The rate of induction of dominant lethals in *Drosophila melanogaster* sperm by X-rays. *J. Genet.* **47**, 1.

Catcheside, D. G. & Lea, D. E. (1945*b*). Dominant lethals and chromosome breaks in ring-*X* chromosomes of *Drosophila melanogaster*. *J. Genet.* **47**, 25.

Demerec, M. & Fano, U. (1944). Frequency of dominant lethals induced by irradiation in sperms of *Drosophila melanogaster*. *Genetics*, **29**, 348.

Fano, U. (1942). Mechanism of induction of gross chromosomal rearrangements in *Drosophila* sperms. *Proc. Nat. Acad. Sci., Wash.*, **29**, 12.

Fisher, R. A. & Yates, F. (1938). *Statistical Tables for Biological, Agricultural and Medical Research*. London: Oliver and Boyd.

Lea, D. E. & Catcheside, D. G. (1942). The mechanism of the induction by radiation of chromosome aberrations in *Tradescantia*. *J. Genet.* **44**, 216.

Lea, D. E. & Catcheside, D. G. (1945). The relation between recessive lethals, dominant lethals, and chromosome aberrations in *Drosophila*. *J. Genet.* **47**, 10.

Muller, H. J. (1940). An analysis of the process of structural change in chromosomes of *Drosophila*. *J. Genet.* **40**, 1.

Muller, H. J. (1941). Induced mutations in *Drosophila*. *Cold Spr. Harb. Symp.* **9**, 151.

Pontecorvo, G. (1942). The problem of dominant lethals. *J. Genet.* **43**, 295.

Sax, K. (1940). An analysis of X-ray induced chromosomal aberrations in *Tradescantia*. *Genetics*, **25**, 41.

Stadler, L. J. (1941). The comparison of ultraviolet and X-ray effects on mutation. *Cold Spr. Harb. Symp.* **9**, 168.

W.P.A. Mathematical Tables (1939–41). New York: Works Projects Administration.