Theoretical and Applied Genetics 44, 31–44 (1974)

A Mathematical Investigation of the Induced Mutation Rate which is Optimum for Genetic Improvement

Part I. Mutagenic Treatment of the Haploid: the Three-Locus Case*

BASIL DIAMANTIS

Department of Genetics and Cell Biology, University of Minnesota, St. Paul, Minnesota (USA)

Summary. Even if there is a high ratio of unfavorable to favorable mutations, there is still a finite probability that a favorable genotype will arise by mutation, if only favorable mutants happen to occur or if the effects of the favorable ones outweigh the effects of the unfavorable ones. The object of this investigation was to determine the mutation rate (termed the optimum mutation rate) that maximizes the probability of a favorable genotype. This was investigated for a diploid plant with pollen treatment followed by self-fertilization to essential homozygosity.

The parameters considered are block (chromosome) number, number of loci per chromosome, ratio of favorable to unfavorable mutants, and amount of recombination (c). The exact ranges in the interval $0 \leq c \leq 0.5$ have been obtained for the optimum mutation rate and the corresponding probability of obtaining an improved genotype.

In later publications the effects of seed treatment and of random mating before self-fertilization are considered. It is found, and will be discussed in a later publication, that (1) seed treatment is better than pollen treatment, and (2) if the number of loci is sufficiently large the optimum dosage of the mutagen may be so small as to make artificial mutagenic treatment undesirable.

1. Introduction

It has been known since the third decade of this century that irradiation increases mutation rate (Muller 1927, 1928; Goodspeed and Olson 1928; Stadler 1928, 1929; and others) and this is firmly established by voluminous data accumulated since then. Experimental data also indicate that the frequency of induced mutation depends on the total dose of mutagen (Stadler 1930; Muller *et al.* 1954; Newcombe 1955; Gustafsson 1963) and hence that within limits mutation rates are subject to control.

Many investigators have been and are using mutagenic agents to produce new alleles that are better than the pre-existing ones so that superior genotypes can be produced (Gustafsson 1947, 1963; Gregory 1955, 1956; Cooper and Gregory 1960; Gustafsson *et al.* 1960; Gaul 1961, 1965; Tavcar 1965; Brock 1965; Frey 1965; Pfeifer 1965; Scossiroli 1965; and many others).

In general, unfavorable mutants will greatly outnumber those that improve the performance, so it is clear that a very high rate of mutation will produce so many unfavorable mutants as to swamp the smaller number of favorable ones. On the other hand, too low a mutation rate will produce no mutants at all in most trials. Clearly, there is an intermediate mutation rate that is optimum in that it maximizes the probability of getting an overall improvement.

The situation considered is that of a self-fertilizing plant. Either seed or pollen is treated with a mutagenic agent and the progeny are continuously self-fertilized until essentially homozygous before testing. The object is to determine the mutation rate that maximizes the probability of obtaining an improved line, either by obtaining only favorable mutations or by having the ratio of favorable to unfavorable such that the effect of the favorable ones outweighs the effect of the harmful ones. A variation of the procedure is to permit random mating for one or more generations before starting self-fertilization.

In this paper I consider self-fertilization following the treatment of pollen. Seed treatment and random mating before self-fertilization are treated in a subsequent paper. The parameters considered are the number of blocks (chromosomes), the number of loci per block, the amount of recombination, and the ratio of favorable to unfavorable mutations.

For the purpose of this study the genotype is visualized as comprised of blocks (a block may be identified with a chromosome) that carry particulate units of inheritance at specific loci. Let m be the number of these blocks, n the number of loci per block and c the probability of genetic recombination between adjacent loci. Further, for any locus, let 0 signify the allele present in the homozygous material prior to mutagenic treatment, let $+$ indicate any mutant allele of 0 that increases the value of the total genotype and let $-$ represent any mutant allele of 0 that decreases the total value of the genotype. Let p_1 be the rate at which the 0 allele mutates to a $+$

---

* This paper was supported in part by Grant Number GM15422 from the National Institutes of Health. Paper 01565 from the Laboratory of Genetics, University of Wisconsin-Madison.

allele, p_2 be the rate at which the 0 allele mutates to a − allele and p_0 be the probability of no mutation. The three probabilities must add to unity.

$$p_1 + p_2 + p_0 = 1 \tag{1}$$

Next let the total mutation rate be

$$p = p_1 + p_2 \tag{2}$$

and the ratio of mutation rates to + and − alleles be

$$k = \frac{p_1}{p_2} \tag{3}$$

Then, p_0, p_1 and p_2 can be expressed in terms of p and k or

$$p_0 = 1 - p;\ p_1 = \frac{k\,p}{1+k};\ p_2 = \frac{p}{1+k}. \tag{4}$$

There is ample experimental evidence that $p_2 > p_1$, i. e., that $k < \frac{1}{2}$, probably very considerably less than one-half.

Two modes of classification of homozygous lines will be employed that are based on the nature of genetic blocks that make up the genotype. A third mode will be considered in a later paper that is based on numbers of 0, + and − genes in the genotype. We shall distinguish four types of genetic blocks as follows:

Type	Number of +, − and 0 alleles in the block		
	$n(+)$	$n(-)$	$n(0)$
a	1	0	$n-1$
b	>1	0	$n - n(+)$
c	≥ 0	≥ 1	$n - n(+) - n(-)$
d	0	0	n

1. For the first mode of classification of lines let m_1 be the number of genetic blocks that are type a or type b, m_2 be the number of blocks that are type c, and $m_3 = m - m_1 - m_2$ be the number of blocks that are type d. Then all the lines for which m_1 and m_2 are the same comprise a class of lines and the probability of a line belonging to such a class will be designated $w(p; c, m_1, m_2, m, n)$.

2. For the second mode of classification let m_1 be the number of genetic blocks of type a and m_2 and $m_3 = m - m_1 - m_2$ be as defined for the first mode. The probability of a line belonging to a specified class of lines will be designated $\omega(p; c, m_1, m_2, m, n)$.

When both + and − alleles are present in the genotype of a line, that genotype may or may not be superior to the original genotype (in which there are only 0 alleles) depending on whether the sum of effects of all + alleles is greater in absolute magnitude than the sum of effects of the − alleles present. Thus the only classes of lines of which all lines are certainly superior are those for which $m_2 = 0$. For this reason these classes will receive special attention in what follows.

2. Derivations

With respect to *a specific block*, the *population* of diploid genotypes that are carried by plants arising immediately from the fertilization of untreated ovules by the treated pollen, will be referred to as the *"initial"* population.

All that follows is based on the assumptions

(a) Genetic segregation between blocks is independent.

(b) The mutational events are independent between loci so that their joint probabilities are the products of their absolute probabilities.

(c) The mutation rates, i. e., the p_1's are equal for all loci and also the p_2's; and the p_0's.

(d) Changes in the amount of radiation or other mutagens cause proportional changes in p_1 and p_2 so that $\frac{p_1}{p_2} = k$ where k is a positive constant.

(e) Mutation has no effect on fitness, i. e., all individuals carrying the treated genetic material have equal reproductive fitness.

(f) The recombination fractions for all adjacent pairs of loci and all genetic blocks are equal, i. e., $c_{ij} = c$ for all i, j; $i = 1, \ldots, m$, $j = 1, \ldots, n-1$ (m and n are as defined in the introduction and genetic interference is absent).

On the basis of *assumption* (a) and the types of genetic blocks defined in the introduction,

mode 1:

$$w(p; c, m_1, m_2, m, n) = \frac{m!}{m_1!\, m_2!\, (m - m_1 - m_2)!}\, z_n^{m_1}\, g_n^{m_2}\, h_n^{m - m_1 - m_2} \tag{5}$$

and mode 2:

$$\omega(p; c, m_1, m_2, m, n) = \frac{m!}{m_1!\, m_2!\, (m - m_1 - m_2)!}\, t_n^{m_1}\, g_n^{m_2}\, h_n^{m - m_1 - m_2} \tag{6}$$

where, in homozygous lines obtained by continuous self-fertilization,

z_n is the probability that any specific genetic block will be of type a or b, i. e., that of n loci one or more will be homozygous for a + allele and the remainder for 0 alleles,

g_n is the probability that any specific genetic block will be of type c, i. e., that of n loci one or more will be homozygous for a − allele,

h_n is the probability that any specific genetic block will be of type d, i. e., that all n loci will be homozygous for 0 alleles, and t_n is the probability that any specific genetic block will be of type a, i. e., that of n loci one will be homozygous for a + allele and $n - 1$ homozygous for 0 alleles.

The practical use of equations (5) and (6) depends on finding explicit expressions for z_n, g_n, h_n and t_n. Let the "initial" population of the diploid genotypes which may be carried by any one block be denoted by $Z_i : i = 1, \ldots, 3^n$, where Z_i is the i^{th} genotype.

Let the population of homozygous genotypes that may be carried in a block be designated by $L_j : j = 1, \ldots, 3^n$ and the probability of the j^{th} homozygous genotype be $P(L_j)$. These individual probabilities will be obtained making use of the following

$$P(L_j) = \sum_i P(Z_i) P(L_j/Z_i) \qquad (7)$$

where $P(Z_i)$ is the probability of a block having the i^{th} genotype in the "initial" population and $P(L_j/Z_i)$ is the probability that continuous self-fertilization will produce homozygous genotype j given the inital genotype, Z_i. Summation of $P(L_j)$ for all genotypes of a particular type, e.g.,

$$\sum_j P(L_{ja}) = \sum_i P(Z_i) \sum_j P(L_{ja}/Z_i) \qquad (8)$$

then provides z, g, h and t.

$$\left.\begin{aligned} z &= \sum_j P(L_{ja}) + \sum_j P(L_{jb}) = P_a + P_b \\ g &= \sum_j P(L_{jc}) = P_c \\ h &= \sum_j P(L_{jd}) = P_d \\ t &= \sum_j P(L_{ja}) = P_a \end{aligned}\right\} \qquad (9)$$

where L_{ja}, L_{jb}, L_{jc} and L_{jd} symbolize homozygous genotypes of the four types explained in the introduction. The four sums in (9) will be referred to alternatively as $P_e : e = a, \ldots, d$. For example $P_a = \sum_j P(L_{ja})$.

In the remainder of this section the probabilities, z, g, h and t, will be derived for various situations specified in terms of n, c. On the basis of these, optimum mutation rates and other results of special interest are obtained.

Number of loci: $n = 2$

$w\,(p;\ c, m_1, m_2, m, 2)$

We proceed first to obtain explicit expressions for z_2, g_2 and h_2. The first step is specification of the Z_i and $P(Z_i)$. Mutagenic treatment of pollen gives rise to nine gametic possibilities (*assumptions* (b) and (c), stated earlier). These with their probabilities are

gamete	++	+0	+−	0+	00	0−	−+	−0	−−
probability	p_1^2	p_1p_0	p_1p_2	p_1p_0	p_0^2	p_0p_2	p_1p_2	p_0p_2	p_2^2

Fertilization of non-treated material (possessing only 0 genes) by the treated pollen yields nine kinds of genotypes, i.e., the population $Z_i : i = 1, \ldots, 3^2$ (given in table 1) which have the same probability distribution as that of the gametes above. Hence

$$P(Z_i) = p_1^x p_0^y p_2^{2-x-y} \quad x = 0, 1, 2;\ y = 0, 1, 2 \quad x + y \leq 2 \qquad (10)$$

where x, y and 2-x-y are, respectively, the numbers of loci with $+$, 0 and $-$ genes in the haploid block contributed to Z_i by the treated pollen.

Selfing begins on the Z_i and after a large number of generations of self-fertilization homozygous genotypes of the following kinds $(L_j : j = 1, , \ldots, 3^2)$ are obtained.

Genotype	*Symbol*	*Genotype*	*Symbol*	*Genotype*	*Symbol*
++/++	L_{1b}	0+/0+	L_{4a}	−+/−+	L_{7c}
+0/+0	L_{2a}	00/00	L_{5d}	−0/−0	L_{8c}
+−/+−	L_{3c}	0−/0−	L_{6c}	−−/−−	L_{9c}

From equation (10) we have the $P(Z_i)$. In order to obtain the P_e by (9) we require the frequencies of the homozygotes produced through selfing the heterozygote AB/ab. These frequencies are (Diamantis 1973)

genotype	AB/AB	Ab/Ab	aB/aB	ab/ab
frequency	$\frac{1}{2(1+2c)}$	$\frac{c}{1+2c}$	$\frac{c}{1+2c}$	$\frac{1}{2(1+2c)}$

On the basis of these, then, the conditional probabilities appearing in table 1 are derived.

The last row of table 1 gives symbolically the frequencies, P_e. From these z_2, g_2 and h_2, by relations (9) and (4), are

$$z_2 = \frac{kp}{1+k}\left(1 - \frac{(2+k)\,p}{2\,(1+k)\,(1+2c)}\right) \qquad (11)$$

$$g_2 = \frac{p}{1+k}\left(1 - \frac{p}{2\,(1+k)\,(1+2c)}\right) \qquad (12)$$

$$h_2 = 1 - p + \frac{p^2}{2\,(1+2c)} \qquad (13)$$

where c is the frequency of recombination between the two loci and

$$z_2 + g_2 + h_2 = 1.$$

The optimum total mutation rate, denoted by $p_w^0(c, m_1, m_2, m, 2)$, or simply by p_w^0 (for the function w), can be obtained as the value of p which maximizes (5). The derivative

$$\frac{d}{dp}\,w\,(p;\ c, m_1, m_2, m, n)$$

$$= \frac{m!\, z_n^{m_1} g_n^{m_2} h_n^{m-m_1-m_2}}{m_1!\, m_2!\, (m - m_1 - m_2)!}\, \frac{m_1}{z_n}\frac{dz_n}{dp} + \frac{m_2}{g_n}\frac{dg_n}{dp} + {} + \frac{(m - m_1 - m_2)}{h_n}\frac{dh_n}{dp}$$

vanishes when

$$\frac{m_1}{z_n}\frac{dz_n}{dp} + \frac{m_2}{g_n}\frac{dg_n}{dp} + \frac{(m - m_1 - m_2)}{h_n}\frac{dh_n}{dp} = 0 \qquad (14)$$

and one of the roots of (14) is optimum p_w^0 in that it maximizes w.

In practical breeding the optimum mutation rate, p_w^0, has utility only for the case $m_2 = 0$. This is because nothing certain can be stated concerning the relative magnitudes of effects of favorable and unfavorable mutations. As a result a line cannot be drawn between "superior" and "inferior" genotypes when $m_2 \neq 0$.

Table 1. *Population Z_i when $n = 2$ loci per block, and the conditional probabilities of the four types a, b, c, and d of homozygous lines, from which their absolute probabilities P_e are obtained*

			$P(L_{2e}/Z_i)$*; $e = a, b, c, d$			
i	Z_i	$P(Z_i)$	L_{2b}	L_{2a}	L_{2c}	L_{2d}
1	$++/00$	p_1^2	$\frac{1}{2(1+2c)}$	$\frac{2c}{1+2c}$	0	$\frac{1}{2(1+2c)}$
2	$+0/00$	p_1p_0	0	$\frac{1}{2}$	0	$\frac{1}{2}$
3	$+-/00$	p_1p_2	0	$\frac{c}{1+2c}$	$\frac{1}{2}$	$\frac{1}{2(1+2c)}$
4	$0+/00$	p_1p_0	0	$\frac{1}{2}$	0	$\frac{1}{2}$
5	$00/00$	p_0^2	0	0	0	1
6	$0-/00$	p_0p_2	0	0	$\frac{1}{2}$	$\frac{1}{2}$
7	$-+/00$	p_1p_2	0	$\frac{c}{1+2c}$	$\frac{1}{2}$	$\frac{1}{2(1+2c)}$
8	$-0/00$	p_0p_2	0	0	$\frac{1}{2}$	$\frac{1}{2}$
9	$--/00$	p_2^2	0	0	$\frac{1+4c}{2(1+2c)}$	$\frac{1}{2(1+2c)}$
Total frequencies			$\sum_j P(L_{jb})$	$\sum_j P(L_{ja})$	$\sum_j P(L_{jc})$	$\sum_j P(L_{jd})$

* The subscript 2 in L_{2e} stands for $n = 2$ loci per genetic block

That is, when $m_2 \neq 0$, usually both favorable and unfavorable mutations are occurring. This will be advantageous in evolution or to the breeder only if the effect of favorable mutants outweighs that of the unfavorable ones. This case will be discussed in a later paper. For the present we shall consider only the case where $m_2 = 0$.

1. When $m_2 = 0$ (i.e., no unfavorable mutations) we get for two loci

$$w\,(p;c,m_1,0,m,2) = \frac{m!}{m_1!\,(m-m_1)!}\,z_2^{m_1}\,h_2^{m-m_1} \qquad (15)$$

and (14) becomes

$$p^3 + a_1\,p^2 + a_2\,p + a_3 = 0 \qquad (16)$$

where

$$a_1 = -\frac{(1+2c)}{m\,(2+k)}\left(m_1 + m\,(4+3k)\right)$$

$$a_2 = \frac{2\,(1+2c)}{m\,(2+k)}\left(m_1\,(2+k) + m\,(1+k)\,(1+2c)\right)$$

$$a_3 = -\frac{2\,m_1\,(1+k)\,(1+2c)^2}{m\,(2+k)}$$

One of the roots of (16) gives the optimum p_w^0 for (15), or

$$p_w^0\,(c,m_1,m,2) = A - \frac{(3\,a_2 - a_1^2)}{9\,A} - \frac{a_1}{3} \qquad (17)$$

with

$$A = \left(-A' + \left(A'^2 + \left(\frac{3\,a_2 - a_1^2}{9}\right)^3\right)^{\frac{1}{2}}\right)^{\frac{1}{3}}$$

such that $p_w^0 < 1$

and

$$A' = \frac{a_3}{2} - \frac{a_1\,a_2}{6} + \frac{a_1^3}{27}$$

Then, substituting (17) for p in (11) and (13), the maximum of (15) is

$$w\,(p^0;c,m_1,0,m,2) = $$
$$= \frac{m!}{m_1!\,(m-m_1)!}\left(z_2\,(p_w^0)\right)^{m_1}\left(h_2\,(p_w^0)\right)^{m-m_1}. \qquad (18)$$

The graphs of (15) for $m = 3, 10, 20$, $m_1 = 1$ and $k = 0.10$ are shown in Figs. 1, 2, and 3.

2. When $m_2 = 0$ and $m - m_1 = 0$ (i.e., *at least one* favorable mutation in *each* block) then $m = m_1$ and we have the special case of [5]

$$w = z_2^m = u\,(p;c,m,2) = $$
$$= \left(\frac{k\,p}{1+k}\left(1 - \frac{(2+k)\,p}{2\,(1+k)\,(1+2c)}\right)\right)^m \qquad (19)$$

by equation (11). Then

$$\frac{d}{dp}\,u\,(p;c,m,2) = 0$$

gives the optimum for the function $u(p;c,m,2)$, or

$$p_u^0(c,2) = \frac{(1+k)\,(1+2c)}{2+k}\ \text{for } c < \frac{1}{2\,(1+k)} \qquad (20)$$

and

$$p_u^0(c,2) = 1 \qquad \text{when } c \geq \frac{1}{2\,(1+k)}.$$

Substituting (20) into (19) yields the maximum

$$u\,(p^0;c,m,2) = \left(\frac{k\,(1+2c)}{2\,(2+k)}\right)^m \text{if } c < \frac{1}{2\,(1+k)} \qquad (21)$$

and

$$u(p^0; c, m, 2) = \left(\frac{k\,(k + 4\,(1 + k)\,c)}{2\,(1 + k)^2\,(1 + 2\,c)}\right)^m \text{if } c \geq \frac{1}{2\,(1 + k)}.$$

The graphs of (19) for $m = 2$ and $k = 0.05$, 0.10 are shown in figs. 6 and 7.

$\omega(p; c, m_1, m_2, m, 2)$

The probabilities g_2 and h_2 are given above and t_2 obtained using (9), (4) and table 1 is

$$t_2 = \frac{k\,p}{1 + k}\left(1 - \frac{p}{1 + 2\,c}\right). \tag{22}$$

The sum of g_2, h_2 and t_2 is less than one.

The mutation rate that maximizes the function ω is given by one of the solutions of $\frac{d\omega}{dp} = 0$ and in particular by a solution of the following parallel of (14)

$$\frac{m_1}{t_n}\frac{dt_n}{dp} + \frac{m_2}{g_n}\frac{dg_n}{dp} + \frac{(m - m_1 - m_2)}{h_n}\frac{dh_n}{dp} = 0 \tag{23}$$

1. When $m_2 = 0$ we obtain the special case of ω

$$\omega\,(p; c, m_1, 0, m, 2) = \frac{m!}{m_1!\,(m - m_1)!}\,t_2^{m_1}\,h_2^{m - m_1} \tag{24}$$

and (23) becomes

$$p^3 + a_1' p^2 + a_2' p + a_3' = 0 \tag{25}$$

where

$$a_1' = -\frac{(1 + 2\,c)}{2\,m}\,(m_1 + 4\,m)$$

$$a_2' = \frac{(1 + 2\,c)}{m}\,\big(2\,m_1 + m\,(1 + 2c)\big)$$

$$a_3' = -\frac{m_1}{m}\,(1 + 2\,c)^2.$$

The optimum p_ω^0 for (24) can be found if the coefficients a_1', a_2' and a_3' are substituted, respectively, for a_1, a_2 and a_3 in (17). Then the maximum of ω can be obtained by replacing p in (24) by p_ω^0.

2. When $m_2 = 0$ and $m - m_1 = 0$ then $m = m_1$ and we have the special case of (6), (i.e., *just one* favorable mutation in *each* block),

$$\omega = t_2^m = v(p; c, m, 2) = \left(\frac{k\,p}{1 + k}\left(1 - \frac{p}{1 + 2\,c}\right)\right)^m \tag{26}$$

by equation (22). Then

$$\frac{d}{dp}\,v(p; c, m, 2) = 0$$

gives the optimum mutation rate for the function $v(p; c, m, 2)$ or

$$p_v^0\,(c, 2) = \frac{1 + 2\,c}{2} \tag{27}$$

Substituting p_v^0 for p into (26) we obtain the maximum of v, or

$$v\,(p^0;\,c,\,m,\,2) = \left(\frac{k\,(1 + 2\,c)}{4\,(1 + k)}\right)^m. \tag{28}$$

The graphs of (26) for $m = 2$ and $k = 0.05$, 0.10 are shown in figs. 8 and 9.

Number of loci: $n = 3$

$w(p; c, m_1, m_2, m, 3)$

With respect to a specific block, fertilization of non-treated, 000/000, ovules by the treated pollen gives rise to the intial population of diploid genotypes $Z_i : i = 1, \ldots, 3^3$ with a probability distribution equal to that of the treated gametes which entered the Z_i, or

$$P(Z_i) = p_1^x p_0^y p_2^{3-x-y} \quad x = 0, \ldots, 3; \quad y = 0, \ldots, 3$$
$$x + y \leq 3 \tag{29}$$

given explicitly in table 2. The exponents x, y and $3 - x - y$ are, respectively, the numbers of loci with +, 0 and − genes in the haploid block contributed to Z_i by the treated pollen.

The aggregate of homozygous genotypes $L_j : j = 1, \ldots, 3^3$ is as follows:

Genotype	*Symbol*	*Genotype*	*Symbol*	*Genotype*	*Symbol*
+++/+++	L_{1b}	0++/0++	L_{10b}	−++/−++	L_{19c}
++0/++0	L_{2b}	0+0/0+0	L_{11a}	−+0/−+0	L_{20c}
++−/++−	L_{3c}	0+−/0+−	L_{12c}	−+−/−+−	L_{21c}
+0+/+0+	L_{4b}	00+/00+	L_{13a}	−0+/−0+	L_{22c}
+00/+00	L_{5a}	000/000	L_{14d}	−00/−00	L_{23c}
+0−/+0−	L_{6c}	00−/00−	L_{15c}	−0−/−0−	L_{24c}
+−+/+−+	L_{7c}	0−+/0−+	L_{16c}	−−+/−−+	L_{25c}
+−0/+−0	L_{8c}	0−0/0−0	L_{17c}	−−0/−−0	L_{26c}
+−−/+−−	L_{9c}	0−−/0−−	L_{18c}	−−−/−−−	L_{27c}

In order to derive the P_e we require the frequencies of the homozygotes produced by selfing from the initial heterozygote *ABD*/*abd*. These are symbolized as follows:

Genotype	*Probab.*	*Genotype*	*Probab.*
ABD/*ABD*	π_1	*aBD*/*aBD*	π_5
ABd/*ABd*	π_2	*aBd*/*aBd*	π_6
AbD/*AbD*	π_3	*abD*/*abD*	π_7
Abd/*Abd*	π_4	*abd*/*abd*	π_8

The frequencies π_i, $i = 1, \ldots, 8$, may be obtained from the frequencies of homozygotes bearing a two locus genotype.

Considering pairs of loci

$$\pi_7 + \pi_8 = \pi_1 + \pi_2 = \pi_1 + \pi_5 = \pi_4 + \pi_8 = \frac{1}{2\,(1 - 2\,c)} \tag{30}$$

hence,

$$\pi_2 = \pi_4 = \pi_5 = \pi_7\,.$$

Also,

$$\pi_1 + \pi_3 = \pi_6 + \pi_8 = \frac{1}{2\,(1 + 4\,c - 4\,c^2)} \tag{31}$$

since the recombination fraction between *loci a* and *d* is

$$c_3 = c_1 + c_2 - 2\,c_1\,c_2 = 2\,c\,(1 - c)\,.$$

Table 2. *Population Z_i for $n = 3$ loci per genetic block and the conditional probabilities of the homozygous lines a, b, c and d, from which their absolute probabilities P_e are obtained*

i	Z_i	$P(Z_i)$	$P(L_{3e}/Z_i)$;* $e = a, b, c, d$; $i = 1, \ldots, 3^3$			
			L_{3b}	L_{3a}	L_{3c}	L_{3d}
1	++ +/000	p_1^3	$\frac{1}{2}$	$\frac{c(2+5c-6c^2)}{(1+2c)(1+4c-4c^2)}$	0	$\frac{4c^3-6c^2+2c+1}{2(1+2c)(1+4c-4c^2)}$
2	++0/000	$p_1^2 p_0$	$\frac{1}{2(1+2c)}$	$\frac{2c}{1+2c}$	0	$\frac{1}{2(1+2c)}$
3	++−/000	$p_1^2 p_2$	$\frac{c(1-c)}{1+4c-4c^2}$	$\frac{c}{1+2c}$	$\frac{1}{2}$	$\frac{4c^3-6c^2+2c+1}{2(1+2c)(1+4c-4c^2)}$
4	+0+/000	$p_1^2 p_0$	$\frac{1}{2(1+4c-4c^2)}$	$\frac{4c(1-c)}{1+4c-4c^2}$	0	$\frac{1}{2(1+4c-4c^2)}$
5	+00/000	$p_1 p_0^2$	0	$\frac{1}{2}$	0	$\frac{1}{2}$
6	+0−/000	$p_1 p_0 p_2$	0	$\frac{2c(1-c)}{1+4c-4c^2}$	$\frac{1}{2}$	$\frac{1}{2(1+4c-4c^2)}$
7	+−+/000	$p_1^2 p_2$	$\frac{(3-2c)c^2}{(1+2c)(1+4c-4c^2)}$	$\frac{2c(1-c)}{1+4c-4c^2}$	$\frac{1}{2}$	$\frac{4c^3-6c^2+2c+1}{2(1+2c)(1+4c-4c^2)}$
8	+−0/000	$p_1 p_0 p_2$	0	$\frac{c}{1+2c}$	$\frac{1}{2}$	$\frac{1}{2(1+2c)}$
9	+−−/000	$p_1 p_2^2$	0	$\frac{c(1-c)}{1+4c-4c^2}$	$\frac{-16c^3+12c^2+8c+1}{2(1+2c)(1+4c-4c^2)}$	$\frac{4c^3-6c^2+2c+1}{2(1+2c)(1+4c-4c^2)}$
10	0++/000	$p_1^2 p_0$	$\frac{1}{2(1+2c)}$	$\frac{2c}{1+2c}$	0	$\frac{1}{2(1+2c)}$
11	0+0/000	$p_1 p_0^2$	0	$\frac{1}{2}$	0	$\frac{1}{2}$
12	0+−/000	$p_1 p_0 p_2$	0	$\frac{c}{1+2c}$	$\frac{1}{2}$	$\frac{1}{2(1+2c)}$
13	00+/000	$p_1 p_0^2$	0	$\frac{1}{2}$	0	$\frac{1}{2}$
14	000/000	p_0^3	0	0	0	1
15	00−/000	$p_0^2 p_2$	0	0	$\frac{1}{2}$	$\frac{1}{2}$
16	0−+/000	$p_1 p_0 p_2$	0	$\frac{c}{1+2c}$	$\frac{1}{2}$	$\frac{1}{2(1+2c)}$
17	0−0/000	$p_0^2 p_2$	0	0	$\frac{1}{2}$	$\frac{1}{2}$
18	0−−/000	$p_0 p_2^2$	0	0	$\frac{1+4c}{2(1+2c)}$	$\frac{1}{2(1+2c)}$
19	−++/000	$p_1^2 p_2$	$\frac{c(1-c)}{1+4c-4c^2}$	$\frac{c}{1+2c}$	$\frac{1}{2}$	$\frac{4c^3-6c^2+2c+1}{2(1+2c)(1+4c-4c^2)}$

20	$-+0/000$	$p_1 p_0 p_2$	0	$\frac{c}{1+2c}$	$\frac{1}{2}$	$\frac{1}{2(1+2c)}$
21	$-+-/000$	$p_1 p_2^2$	0	$\frac{(3-2c)c^2}{(1+2c)(1+4c-4c^2)}$	$\frac{-8c^2+8c+1}{2(1+4c-4c^2)}$	$\frac{4c^3-6c^2+2c+1}{2(1+2c)(1+4c-4c^2)}$
22	$-0+/000$	$p_1 p_0 p_2$	0	$\frac{2c(1-c)}{1+4c-4c^2}$	$\frac{1}{2}$	$\frac{1}{2(1+4c-4c^2)}$
23	$-00/000$	$p_0^2 p_2$	0	0	$\frac{1}{2}$	$\frac{1}{2}$
24	$-0-/000$	$p_0 p_2^2$	0	0	$\frac{-8c^2+8c+1}{2(1+4c-4c^2)}$	$\frac{1}{2(1+4c-4c^2)}$
25	$--+/000$	$p_1 p_2^2$	0	$\frac{c(1-c)}{1+4c-4c^2}$	$\frac{1+4c}{2(1+2c)}$	$\frac{4c^3-6c^2+2c+1}{2(1+2c)(1+4c-4c^2)}$
26	$--0/000$	$p_0 p_2^2$	0	0	$\frac{1+4c}{2(1+2c)}$	$\frac{1}{2(1+2c)}$
27	$---/000$	p_2^3	0	0	$\frac{-20c^3+14c^2+10c+1}{2(1+2c)(1+4c-4c^2)}$	$\frac{4c^3-6c^2+2c+1}{2(1+2c)(1+4c-4c^2)}$
Total frequencies			$\sum_j P(L_{jb})$	$\sum_j P(L_{ja})$	$\sum_j P(L_{jc})$	$\sum_j P(L_{jd})$

* The subscript 3 in L_{3e} stands for number of loci $n = 3$ per genetic block.

Therefore

$$\pi_3 = \pi_6 .$$

Finally,

$$\pi_2 + \pi_6 = \pi_2 + \pi_3 = \frac{c}{1+2c} . \qquad (32)$$

Thence

$$\pi_1 = \pi_8 = \frac{4c^3 - 6c^2 + 2c + 1}{2(1+2c)(1+4c-4c^2)}$$

from (30), (31) and (32)

$$\pi_2 = \pi_4 = \pi_5 = \pi_7 = \frac{c(1-c)}{1+4c-4c^2}$$

from above and (30)

$$\pi_3 = \pi_6 = \frac{(3-2c)c^2}{(1+2c)(1+4c-4c^2)}$$

from above and (32)

where

$$\sum_{i=1}^{8} \pi_i = 2(\pi_1 + 2\pi_2 + \pi_3) = 1$$

On the basis of these frequencies the conditional probabilities $P(L_{3e}/Z_i)$ are derived and table 2 is constructed from which the $P_e = \sum_j P(L_{je})$ are obtained.

Then the probabilities in the w-function, by relations (4), (9), the frequencies π_i above and the fact that $\pi_1 + 2\pi_2 + \pi_3 = 1/2$, are

$$z_3 = \frac{k}{2(1+k)} \times$$
$$\times\, p\left(3 - \frac{(2+k)(1+4\pi_1)}{(1+k)} p + \frac{2(3+3k+k^2)\pi_1}{(1+k)^2} p^2\right) \qquad (33)$$

$$g_3 = \frac{1}{2(1+k)} p\left(3 - \frac{(1+4\pi_1)}{(1+k)} p + \frac{2\pi_1}{(1+k)^2} p^2\right) \qquad (34)$$

$$h_3 = 1 - \frac{3}{2} p + \frac{(1+4\pi_1)}{2} p^2 - \pi_1 p^3 \qquad (35)$$

where π_1 is one half of the frequency of the parental type homozygotes obtained by selfing the triple heterozygote, and

$$z_3 + g_3 + h_3 = 1 .$$

Next the optimum mutation rate $p_w^0(c, m_1, m_2, m, 3)$ for the w-function may be found by solution of (14) after substitution of z_3, g_3 and h_3 for z_n, g_n and h_n.

When $m_2 = 0$, $m_1 \neq m$ the equation to be solved is of degree greater than four. General solutions for such equations are not known but they can be solved by various methods of successive approximation.

The graphs of the function w for $n = 3$; $m_1 = 1$; $m_2 = 0$; $m = 3$, 10, 20; $k = 0.10$ are shown in figs. 10, 11 and 12.

Now, when $m_1 = m$ we have the special case of $w(p; c, m_1, m_2, m, 3)$

$$w(p; c, m, 3) = z_3^m = u(p; c, m, 3) = \\ = \left(\frac{k\,p}{2\,(1+k)}\left(3 - \frac{(2+k)\,(1+4\,\pi_1)}{(1+k)}\,p + \right.\right. \\ \left.\left. + \frac{2\,(3+3\,k+k^2)\,\pi_1}{(1+k)^2}\,p^2\right)\right)^m \quad (36)$$

by equation (33). Then

$$\frac{d}{dp}\,u(p; c, m, 3) = 0$$

gives the optimum

$$p_u^0(c, 3) = \\ = \frac{(1+k)}{6\,(3+3\,k+k^2)\,\pi_1}\left((2+k)\,(1+4\,\pi_1) - (d_1)^{1/2}\right) \quad (37)$$

where

$$d_1 = (2+k)^2\,(1+4\,\pi_1)^2 - 18\,(3+3\,k+k^2)\,\pi_1\,.$$

Substitution of (37) into (36) gives the following maximum of (36)

$$u(p^0; c, m, 3) = \\ = \left(\frac{k}{3\,(6\,(3+3\,k+k^2)\,\pi_1)^2}\left((d_1)^{3/2} - (2+k)\,(1+4\pi_1)\,d_2\right)\right)^m \quad (38)$$

with

$$d_2 = (2+k)^2\,(1+4\,\pi_1)^2 - 27\,(3+3\,k+k^2)\,\pi_1$$

$\omega(p; c, m_1, m_2, m, 3)$

The probabilities g_3 and h_3 are given above, (34) and (35), and t_3; obtained by using (9), (4) and table 2, is

$$t_3 = \frac{k\,p}{2\,(1+k)}\,(3 - 2\,(1+4\,\pi_1)\,p + 6\,\pi_1\,p^2) \quad (39)$$

with π_1 as in (34) and

$$t_3 + g_3 + h_3 < 1$$

The mutation rate, $p_\omega^0(c, m_1, m_2, m, 3)$ that maximizes $\omega(p; c, m_1, m_2, m, 3)$ is obtained by solution of (23) after substitution of t_3, g_3 and h_3 for t_n, g_n and h_n.

When $m_2 = 0$, $m_1 \neq m$ the solution must be obtained by successive approximation. The result then can be substituted into the ω-function to provide the maximum of that function.

When $m_1 = m$ the function ω becomes

$$\omega(p; c, m, 3) = t_3^m = v(p; c, m, 3) = \\ = \left(\frac{k\,p}{2\,(1+k)}\,(3 - 2\,(1+4\,\pi_1)\,p + 6\,\pi_1\,p^2)\right)^m \quad (40)$$

by (39). Then

$$\frac{d}{dp}\,v(p; c, m, 3) = 0$$

yields the optimum mutation rate for the function $v(p; c, m, 3)$, or

$$p_v^0(c, 3) = \frac{1}{9\,\pi_1}\left(1 + 4\,\pi_1 - (d_3)^{1/2}\right) \quad (41)$$

where

$$d_3 = (1+4\,\pi_1)^2 - \frac{27}{2}\,\pi_1\,.$$

Substituting $p_v^0(c, 3)$ into (40), the maximum of v is

$$v(p^0; c, m, 3) \\ = \left(\frac{k}{3\,(1+k)\,(9\,\pi_1)^2}\left(2(d_3)^{3/2} - (1+4\,\pi_1)\,d_4\right)\right)^m \quad (42)$$

with

$$d_4 = 2\,(1+4\,\pi_1)^2 - \frac{81}{2}\,\pi_1$$

3. Applications and Discussion

In this section the behavior, with respect to the parameters involved, of the functions derived in the previous section will be examined. Specific values will be assigned to these parameters so that their effects may be studied on the optimum p^0 and its corresponding maximum of the probability functions considered. The parameter m_2 will invariably be given the value zero.

$w(p; c, m_1, 0, m, 2)$

The w cases chosen here are the ones where $m_1 = 1$ and $m = 3, 10, 20$. The reason for selecting these particular values is that the breeder would like to have a pure line whose m-block genome contains, at the least, a single block with one or more + but no−genes, the other m-1 blocks carrying unmutated loci.

These functions have been plotted and their graphs appear in fig. 1, 2 and 3. The number k is given the value 0.10 and c the values 0, 0.25, 0.5.

Fig. 1 shows, for example, that for an organism whose genome is made up of three chromosomes, each carrying two loci, the optimum mutation rate is $p_w^0 \simeq 0.3$ and with this rate there is a probability between 0.03 and 0.04 of obtaining such an organism with one chromosome having at least one favorable mutation in the two homozygous loci and two chromosomes with no mutations when the overall ratio of favorable to unfavorable mutations is 0.1. In fig. 3, representing a genome made up of 20 chromosomes, the optimum mutation rate is $p_w^0 \simeq .05$ and for this there is a probability between 0.0335 and 0.0339 of obtaining a plant with one chromosome having at least one favorable mutation in the two homozygous loci and nineteen chromosomes with no mutations.

We observe in these graphs that as the chance for genetic recombination increases, the maxima of w ascend within all three values of m. However, as m increases the maximum of w, $w(p^0)$, rises within $c = 0$ and it decreases when $c = 0.25$, 0.5. As for

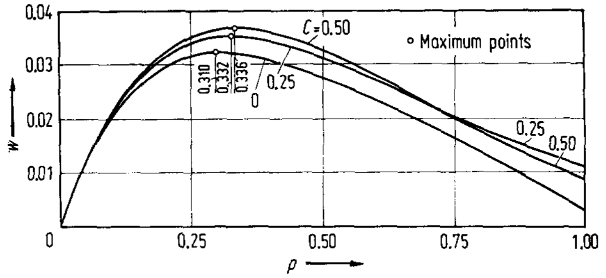

Fig. 1. The graph of the probability function $w(p; c, m_1, 0, m, n)$ of obtaining a homozygous line with a genome consisting of $m = 3$ blocks (each having $n = 2$ loci) of which $m_1 = 1$ block has *at least* one favorable mutation and the remaining $m - m_1 - m_2 = 2$ blocks have no mutations; $m_2 = 0$, i.e., no block has unfavorable mutations. The overall ratio of favorable to unfavorable mutations is $k = 0.1$. The three curves pertain to the three values of the recombination fraction $c = 0.00, 0.25, 0.50$

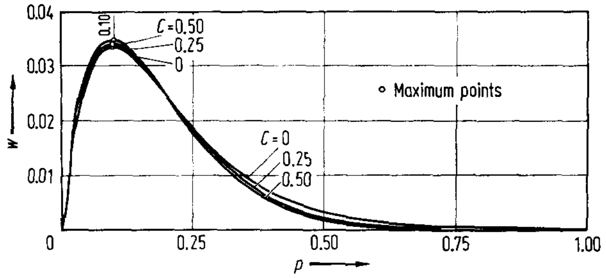

Fig. 2. The graph of the probability function $w(p; c, m_1, 0, m, n)$ of obtaining a homozygous line with a genome containing $m = 10$ genetic blocks (each having $n = 2$ loci) of which $m_1 = 1$ block has *at least* one favorable mutation and the remaining $m - m_1 - m_2 = 9$ blocks have no mutations; $m_2 = 0$, i.e., no block has unfavorable mutations. The overall ratio of favorable to unfavorable mutations is $k = 0.1$

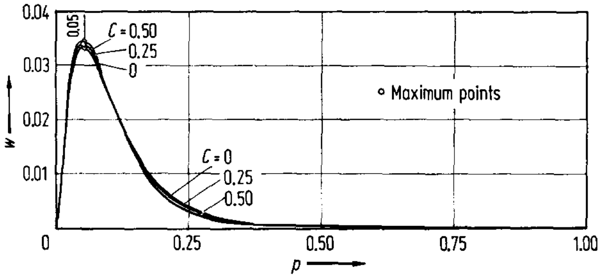

Fig. 3. The graph of the probability function $w(p; c, m_1, 0, m, n)$ of obtaining a homozygous line with a genome comprised of $m = 20$ genetic blocks (each having $n = 2$ loci) of which $m_1 = 1$ block has *at least* one favorable mutation and the remaining $m - m_1 - m_2 = 19$ blocks have no mutations; $m_2 = 0$, i.e., no block has unfavorable mutations. The overall ratio of favorable to unfavorable mutations is $k = 0.1$

the optimum mutation rate, p_w^0, it grows larger with c for $m = 3$, 10 and within $m = 20$ it rises in the interval $0 \leq c \leq 0.25$ but at $c = 0.5$ it falls short of its value at $c = 0.25$. As m increases, however, p_w^0 decreases within each c by roughly two-thirds from $m = 3$ to $m = 10$ and by about one-half from $m = 10$ to $m = 20$ for all c. These facts may also be observed in table 4 at the value of $k = 0.10$.

The explanation of the behavior, just described, of p_w^0 and $w(p^0)$ may be sought in the properties of the components of the w-function.

We observe, from fig. 5, that $h_2 = h(p; c, 1, 2)$ is larger at smaller values of p and c and its maximum occurs at $p = 0$ for all c ($h(p; c, 1, 2)$ does not vanish at $p = 1$ because of the presence of 0 genes coming from the non-treated ova). Whereas, from fig. 4, $z_2 = z(p; c, 1, 2)$ is larger at larger values of p and c so that its maximum with the corresponding optimum p, increases with c.

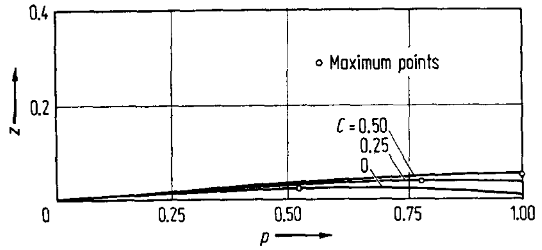

Fig. 4. The component $z_n^{m_1}$ of the w-function where the number of loci $n = 2$ per genetic block. In this example there is only one block, $m_1 = 1$, which is of type a or b, i.e., bearing *at least* one favorable mutation. The components of the function w are the functions $z_n^{m_1}$, $g_n^{m_2}$ and $h_n^{m-m_1-m_2}$ as defined in the text but for this article blocks with unfavorable mutations are not considered so that $m_2 = 0$ and $g_n^{m_2} = 1$

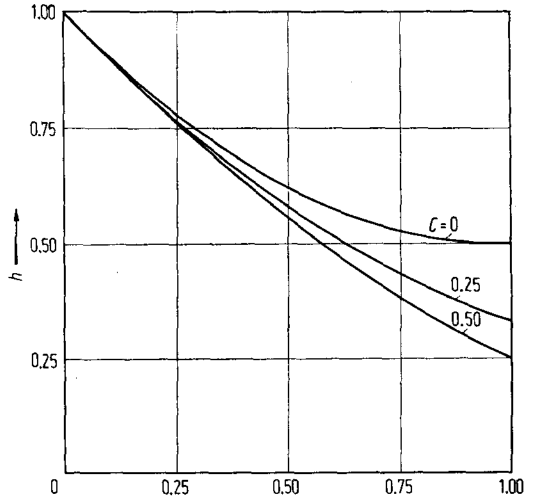

Fig. 5. The component $h_n^{m-m_1-m_2}$ of the w-function where the number of loci $n = 2$ per genetic block. In this graph the exponent of the function h is $m - m_1 - m_2 = 1$, i.e., there is only one block of type d, which is homozygous for unmutated genes. As $m_2 = 0$ the other component of w is the function z represented in fig. 4

First, the function $w(p; c, 1, 0, 3, 2)$, fig. 1, is the product of z_2, the second power of h_2 and a binomial coefficient. This multiplication of the two opposing components of w shifts the maximum $w(p^0)$ and the optimum p_w^0 toward intermediate values, while the influence on them of z_2 with respect to c remains predominant. We think of p_w^0 as the inverse function of $w(p^0)$.

In the function $w(p; c, 1, 0, 10, 2)$ fig. 2, h_2 is represented nine times as much as z_2 so that its influence on w has been magnified:

(a) since h_2 is smaller, whereas z_2 is larger, for larger c, the values of $w(p^0)$ in the interval $0 \leq c \leq 0.5$, have been brought closer together than in the case where $m = 3$; the maximum $w(p^0; c, 1, 0, 10, 2)$ has been shifted toward the origin of the ordinate relative to $w(p^0; c, 1, 0, 3, 2)$ when $c = 0.25$, 0.50 and away from it when $c = 0$,

(b) this shift of $w(p^0)$ is associated with a simultaneous movement of p_w^0 toward the origin of the abscissa (we think of p_w^0 as the inverse function of $w(p^0)$), bringing $p_w^0(c, 1, 0, 10, 2)$ much closer to this point, for all values of c, than $p_w^0(c, 1, 0, 3, 2)$; $p_w^0(c, 1, 0, 10, 2) \simeq \frac{1}{3} p_w^0(c, 1, 0, 3, 2)$.

Finally, in $w(p; c, 1, 0, 20, 2)$ the influence of h_2 has been further enhanced, being raised to the 19th power, thereby narrowing more (than in the case where $m = 10$) the range of $w(p^0)$ in the interval $0 \leq c \leq 0.5$. At the same time p_w^0 falls approximately one-half of its value at $m = 10$, for all c.

Table 3 shows the widths of the ranges of p_w^0 and $w(p^0)$ in the interval $0 \leq c \leq 0.5$.

Table 3. *The widths of the ranges of the optimum mutation rate p_w^0 and its corresponding maximum probability $w(p^0)$ in the interval $0 \leq c \leq 0.5$ for $n = 2$, $k = 0.10$ and $m = 3, 10, 20$. Note the effect of m on these widths, i. e., on the role linkage*

m	$p_w^0(0.5, 1, 0, m, 2) - p_w^0(0, 1, 0, m, 2)$	$w(p^0; 0.5, 1, 0, m, 2) - w(p^0; 0, 1, 0, m, 2)$
3	0.02634	0.00444
10	0.00030	0.00090
20	0,000008*	0.00041

* This entry is $p_w^0(0.25, 1, 0, 20, 2) - p_w^0(0, 1, 0, 20, 2)$; the value of p_w^0 at $c = 0.25$ is the largest of the three values of p_w^0 at $c = 0, 0.25, 0.5$.

These differences fall off rapidly as m becomes larger. When, say, $m \geq 10$ linkage is only very slightly effective in pulling apart the curves of $w(p; c, 1, 0, m, 2)$; in other words, for all practical purposes, neither p_w^0 nor $w(p^0)$ depend on c when $m \geq 10$.

The meaning of the previous findings is that as the number of blocks bearing no mutations increases, in a genome having only one block with + and no − genes, we require a smaller total mutation rate p_w^0 to obtain the pertinent pure line with maximum probability, and that p_w^0 may actually be very small when m is large and may be considered as independent of c.

The joint effects of k, c and m on $p_w^0(c, 1, 0, m, 2)$ and $w(p^0; c, 1, 0, m, 2)$ are shown in table 4. Dominant features are the impact of m on p_w^0 and k on $w(p^0)$.

The special case of w where $m_1 = m$ constitutes the family of functions $u(p; c, m, n)$. Some members of this family for $n = 2$, $m = 2$ and $k = 0.05$, 0.10 appear in fig. 6 and 7.

It is clear from these graphs that as linkage becomes tighter the probability of the desired pure lines becomes smaller and it is the smallest when linkage is complete. Conversely, the optimum $p_u^0(c, 2)$ and its associated maximum $u(p^0; c, 2, 2)$ increase with c. As the probability of mutation becomes larger and linkage is looser there is more opportunity for recombination and this leads to the upward trend in $u(p^0; c, 2, 2)$. In the interval $\frac{1}{2(1+k)} \leq c \leq \frac{1}{2}$ the values of this maximum are associated invariably with one value of the optimum p^0, i.e., $p_u^0 = 1$. This means that, for $n = 2$, when genetic recombination is free or near to it, the total mutation rate and correspondingly the dose of the mutagenic agent have to be the largest in order to obtain the desired pure lines, pertinent to the u-function, with maximum probability.

The effects of the value k on p_u^0 and $u(p^0)$ may be seen in table 5. Comparing the items in columns 2 and 11 of this table we note that for such an increase

Table 4. *The joint effects of linkage, k and m on the optimum mutation rate p_w^0 and its corresponding maximum $w(p^0)$ of the probability function w. Note the impact of m on p_w^0 and of k on $w(p^0)$. Here $n = 2$, $m_1 = 1$, $m_2 = 0$*

m	c	$k = 0.01$		$k = 0.10$		$k = 0.50$	
		$p_w^0(c, 1, 0, m, 2)$	$w(p^0; c, 1, 0, m, 2)$	$p_w^0(c, 1, 0, m, 2)$	$w(p^0; c, 1, 0, m, 2)$	$p_w^0(c, 1, 0, m, 2)$	$w(p^0; c, 1, 0, m, 2)$
3	0.00	0.302947	0.003470	0.310030	0.032426	0.333333	0.125572
	0.25	0.327372	0.003820	0.331560	0.035469	0.344406	0.134578
	0.50	0.333654	0.003983	0.336370	0.036868	0.329694	0.138356
10	0.00	0.099451	0.003631	0.099940	0.033487	0.101430	0.124443
	0.25	0.099909	0.003703	0.100210	0.034095	0.101104	0.126100
	0.50	0.100014	0.003737	0.100240	0.034387	0.083320	0.124634
20	0.00	0.049945	0.003640	0.050057	0.033494	0.050392	0.123597
	0.25	0.0499938	0.003673	0.050065	0.033771	0.047713	0.124177
	0.50	0.0499943	0.003689	0.050059	0.033907	0.032235	0.114342

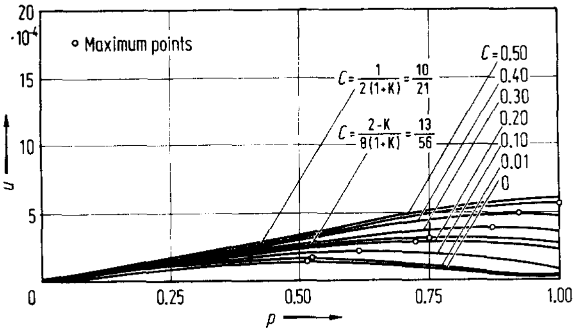

Fig. 6. The special case of w where $m_1 = m$, $m_2 = 0$ and $m - m_1 - m_2 = 0$, constitutes the family of functions $u(p;\, c, m, n)$. Here $n = 2$, $m = 2$. Hence $u(p;\, c, 2, 2)$ is the probability of obtaining a homozygous line whose genome is made up of $m = 2$ genetic blocks, each carrying $n = 2$ loci, and where each block has at *least one* favorable and *no* unfavorable mutations when the overall ratio of favorable to unfavorable mutations is $k = 0.05$. The nine curves refer to the nine values of the recombination fraction shown in the graph

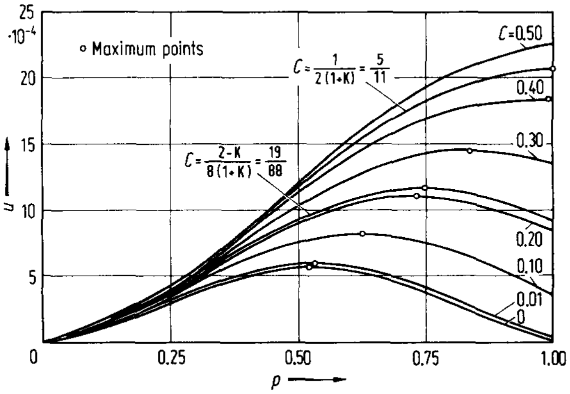

Fig. 7. The function $u(p;\, c, 2, 2)$ of fig. 6 when the ratio of favorable to unfavorable mutations is $k = 0.10$. The nine curves refer to the nine values of c shown in the graph

in k as indicated by the headings of these two columns, the change in p_u^0 is relatively small. However, the corresponding change in the maximum $u(p^0)$ is quite substantial as a contrast of the values in columns 3 and 12 shows. The enhancing effect on the whole function u of doubling k, may be seen by contrasting fig. 6 and 7.

Finally, with respect to m, it is noted from equations (20) and (21) that the optimum p_u^0 is independent of m, whereas its associated maximum $u(p^0)$ decreases rapidly, i.e., exponentially as m becomes larger. The full meaning of this will be discussed in Part II of this series of papers.

$\omega(p;\, c, m_1, 0, m, 2)$

Of this function only the special case where $m_1 = m$ will be examined. This is the function $v(p;\, c, m, 2)$. Curves of the latter from $m = 2$ appear in fig. 8 and 9 for values of k, 0.05 and 0.10 respectively.

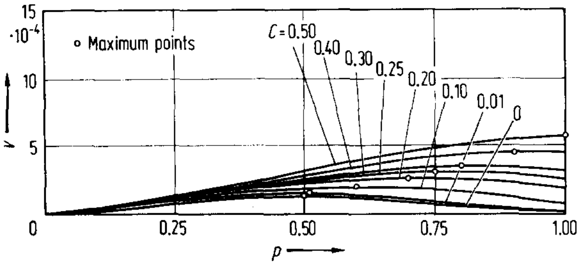

Fig. 8. The special case of ω where $m_1 = m$, $m_2 = 0$ and $m - m_1 - m_2 = 0$, constitutes the family of functions $v(p;\, c, m, n)$. Here $n = 2$, $m = 2$. Hence $v(p;\, c, 2, 2)$ is the probability of recovering a homozygous organism whose genome is made up of $m = 2$ genetic blocks, each bearing $n = 2$ loci, and where each block has *just one* favorable and *no* unfavorable mutations when the overall ratio of favorable to unfavorable mutations is $k = 0.05$. The eight curves refer to the eight values of the recombination fraction shown in the graph

Comparing graphs in fig. 6 and 7 on one hand with those of fig. 8 and 9 on the other it is noted that the u-functions are (as they should be) slightly larger than the v-functions. The recombination fraction c affects the latter in much the same way as the former and the remarks made above for u, p_u^0 and $u(p^0)$ with respect to c apply also to v, p_v^0 $v(p^0)$. However, the

Table 5. *The effects of linkage and the value of k on the optimum mutation rate p_u^0 and its corresponding maximum $u(p^0)$ of the probability function $u\,(p;c,m,n)$. Here $m = 2$, $n = 2$*

c	$k = 0.01$		$k = 0.05$		$k = 0.10$		c	$k = 0.30$		$k = 0.50$	
	p_u^0	$u(p^0)$	p_u^0	$u(p^0)$	p_u^0	$u(p^0)$		p_u^0	$u(p^0)$	p_u^0	$u(p^0)$
0.00	0.503	.0000062	0.512	.000149	0.524	.00057	0.00	0.565	.0043	0.600	.0100
0.01	0.513	.0000064	0.522	.000155	0.534	.00059	0.01	0.577	.0044	0.612	.0104
0.10	0.603	.0000089	0.615	.000214	0.629	.00082	0.10	0.678	.0061	0.720	.0144
0.20	0.704	.0000121	0.717	.000291	0.733	.00111	$\frac{2-k}{8\,(1+k)}$	0.750	.0075	0.750	.0156
$\frac{2-k}{8\,(1+k)}$	0.750	.0000138	0.750	.000319	0.750	.00116	0.20	0.791	.0083	0.840	.0196
0.30	0.804	.0000158	0.820	.000381	0.838	.00145	0.30	0.904	.0109	0.960	.0256
0.40	0.905	.0000201	0.922	.000482	0.943	.00184	$\frac{1}{2\,(1+k)}$	1.000	.0133	1.000	.0278
$\frac{1}{2\,(1+k)}$	1.000	.0000245	1.000	.000567	1.000	.00207	0.40	1.000	.0138	1.000	.0320
0.50	1.000	.0000248	1.000	.000594	1.000	.00226	0.50	1.000	.0166	1.000	.0378

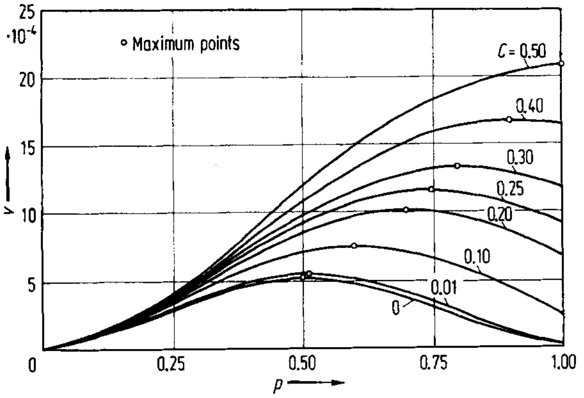

Fig. 9. The function $v(p; c, 2, 2)$ of fig. 8 when the ratio of favorable to unfavorable mutations is $k = 0{,}10$. The eight curves refer to the eight values of c shown in the graph

families of functions u and v differ in the interval $\frac{1}{2(1+k)} \le c \le \frac{1}{2}$ where p_u^0 equals to unity while p_v^0 varies.

A more important difference between u and v is that p_u^0 depends on k whereas p_v^0 does not (compare formulae (20) with (27)). While, however, the optimum mutation rate p_v^0 is unaffected by k, the maximum $v(p^0)$, achieved by that rate, is larger when the value of k is bigger as a contrast of fig. 8 and 9 shows. The significance of this is that in order to obtain the maximum of the u-function, the optimum dose of the mutagen has to be increased as k becomes larger, whereas with the v-function this dose need not be altered if it happens that the value of k changes (provided that in no way k enters the relationship between optimum mutation rate and dose).

Looking at expressions (21) and (28) it can be deduced that the pattern of change of $v(p^0)$ with respect to k is similar to that of $u(p^0)$ presented in table 5. This may be verified by making a contrast of the comparison between fig. 6 and 7 with the comparison between fig. 8 and 9.

From equations (27) and (28) it is noted that (as with the function u) the maximum $v(p^0)$ decreases exponentially as m increases while the optimum p_v^0 is independent of m.

$w(p; c, m_1, 0, m, 3)$

The graphical method was employed to obtain the optimum $p_w^0(c, m_1, 0, m, 3)$ and the associated maximum $w(p^0; c, m_1, 0, m, 3)$, shown in table 6, for the functions represented in fig. 10, 11 and 12. We are dealing with $w(p; c, 1, 0, m, 3)$ for $k = 0.10$ and $m = 3$, 10, 20, that is, with the same functions as in the case where n was equal to 2.

From table 6 it is noted that the effects of changing c within each m and of altering m within each c on $p_w^0(c, 1, 0, m, 3)$ and on $w(p^0; c, 1, 0, m, 3)$ are almost the same as on $p_w^0(c, 1, 0, m, 2)$ and on $w(p^0; c, 1, 0, m, 2)$ discussed above. Since the curves in fig. 10, 11 and 12

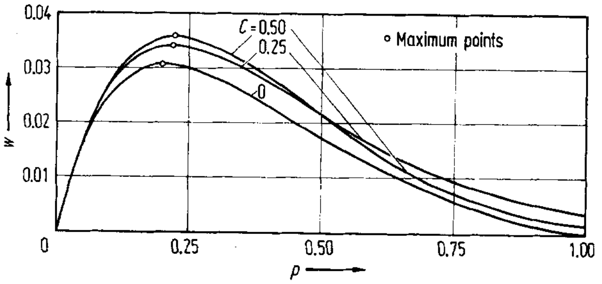

Fig. 10. The graph of the probability function $w(p; c, m_1, 0, m, n)$ of obtaining a homozygous line where each genetic block has $n = 3$ loci but otherwise is the same as in fig. 1

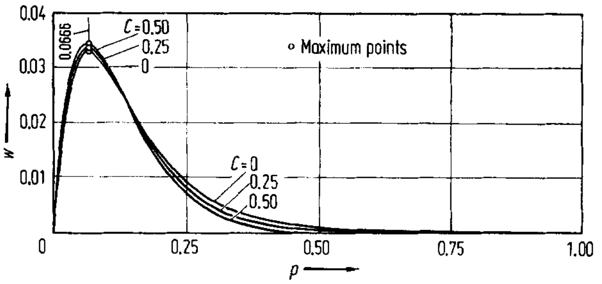

Fig. 11. The graph of the probability function $w(p; c, m_1, 0, m, n)$ of obtaining a homozygous line where each genetic block has $n = 3$ loci but otherwise it is the same as in fig. 2

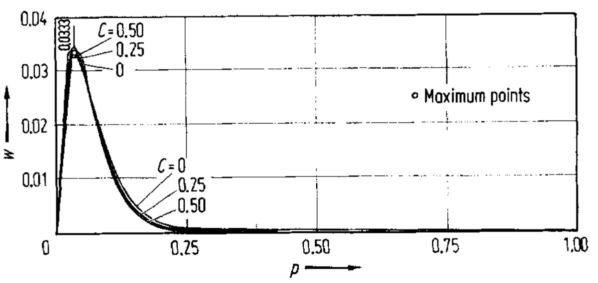

Fig. 12. The graph of the probability function $w(p; c, m_1, 0, m, n)$ of obtaining a homozygous line where each genetic block has $n = 3$ loci but otherwise it is the same as in fig. 3

Table 6. *The effects of linkage and of m on the optimum mutation rate p_w^0 and its corresponding maximum $w(p^0)$ of the probability function w. Note the impact of m on p_w^0 as in the case $n = 2$ (table 4). Here $k = 0.10$, $n = 3$, $m_1 = 1$, $m_2 = 0$*

m	c	$p_w^0(c, 1, 0, m, 3)$	$w(p^0; c, 1, 0, m, 3)$
3	0.00	0.2040	0.0306
	0.25	0.2230	0.0345
	0.50	0.2250	0.0358
10	0.00	0.0667	0.0330
	0.25	0.0669	0.0338
	0.50	0.0669	0.0341
20	0.00	0.0334	0.0332
	0.25	0.0334	0.0336
	0.50	0.0334	0.0338

The values of p_w^0 and $w(p^0)$ are rounded to the 4th decimal place.

of $w(p;\, c, 1, 0, m, 3)$ result from multiplying $z_3 = z(p;\, c, 1, 3)$ with powers of $h_3 = h(p;\, c, 1, 3)$ and a binomial coefficient, the effects of c and m on p_w^0 and $w(p^0)$ for $n = 3$, may be explained along the same lines as for the function $w(p;\, c, 1, 0, m, 2)$, with $n = 2$, outlined above.

The effect of the change in m on the widths of the ranges of $p_w^0(c, 1, 0, m, 3)$ and $w(p^0;\, c, 1, 0, m, 3)$, in the interval $0 \leq c \leq 0.5$, is seen from table 7 to be much the same as in the case where $n = 2$ described above. However, within each m there is a widening of the range of $w(p^0)$ when n goes from 2 to 3, as a comparison of tables 3 and 7 shows, which is contrary to the behavior of $u(p^0)$ and $v(p^0)$ whose ranges, in the interval $0 \leq c \leq 0.5$, become narrower as n increases (compare ranges from table 5 with those from table 8 of $u(p^0)$ for each k). This widening of the range of $w(p^0)$ does not necessarily imply a consistent upward trend of the width of that range as n increases beyond 3.

Table 7. *The widths of the ranges of the optimum mutation rate p_w^0 and its corresponding maximum $w(p^0)$ in the interval $0 \leq c \leq 0.5$ for $n = 3$, $k = 0.10$ and $m = 3, 10, 20$. Note the effect of m on these widths, i. e. on the role of linkage*

m	$p_w^0(0.5, 1, 0, m, 3) - p_w^0(0, 1, 0, m, 3)$	$w(p^0;\, 0.5, 1, 0, m, 3) - w(p^0;\, 0, 1, 0, m, 3)$
3	0.0210	0.0052
10	0.0003	0.0012
20	0.00001*	0.00054

* This entry is $p_w^0(0.25, 1, 0, 20, 3) - p_w^0(0, 1, 0, 20, 3)$; when $m = 20$, p_w^0 has the largest value at $c = 0.25$, the other p_w^0 values considered are at $c = 0$ and $c = 0.5$.

The effects of increasing n from $n = 2$ to $n = 3$ on $p_w^0(c, 1, 0, m, n,)$ and $w(p^0;\, c, 1, 0, m, n)$ will now be considered.

A comparison of graphs in fig. 10, 11 and 12 with those in fig. 1, 2, and 3, respectively, reveals that this increase in the value of n has moved p_w^0 and also $w(p^0)$ toward the origin of the co-ordinate axes, except that $w(p^0)$ has moved away from this point at $c = 0$ and $m = 10, 20$; the shift of p_w^0 toward the origin of the abscissa being greater relative to that of $w(p^0)$ along the ordinate. This decrease in p_w^0, relative to its value at $n = 2$, is on the average approximately 0.33 for each m. (This figure is obtained by subtracting the entries of the 3rd column of table 6 from the corresponding entries of the 5th column of table 4, dividing these differences by the entries of the 5th column of table 4, and averaging over all c for each m.) This is a substantial decrease and it might mean that the total optimum mutation rate $p_w^0(c, 1, 0, m, n)$ is indeed small even when the number of loci per genetic block is not very large, say 200 . In fact it has been found (and will be shown in Part II of this series) that the optima p_u^0 for the function u and p_v^0 for the function v decrease as n increases beyond 3, in the manner just suggested for p_w^0.

The special case of $w(p;\, c, m_1, 0, m, 3)$ for which $m_1 = m$ is the family of functions $u(p;\, c, m, 3)$. It should be noted, as in the case where $n = 2$, that the tighter the linkage the smaller the u-function becomes, and thus its maximum, and also the corresponding optimum p^0. This is shown, for different values of k, in table 8 in which also the effect of the value of k on p_u^0 and $u(p^0)$ is brought out. A contrast of corresponding entries of columns 2 and 10 of table 8 shows that, for the increase in k indicated by the difference of the values at the headings of these columns, the differences in p_u^0 relative to its values at $k = 0.01$, are rather small while the relative differences in $u(p^0)$, i.e., those between corresponding items of columns 3 and 11, relative to the entries of column 3, are much larger. As pointed out earlier this is also the case when $n = 2$.

Now, for any specific value of c and k the effect on p_u^0 and on $u(p^0)$ of increasing n from $n = 2$ to $n = 3$ may be brought out by comparing tables 5 and 8. This increase in n has resulted in a shift of the maximum points toward the origin of the co-ordinate axes, and the shift is greater toward the origin of the abscissa than it is along the ordinate. This would suggest that when n is large the optimum p^0 must be indeed small. This in fact has been found to be so for the functions $u(p;\, c = 0, m, n)$ and $u\left(p;\, c = \frac{1}{2}, m, n\right)$, for any m and n, and will be shown in Part II of this series.

Further, the effect of the increase of the number of loci on the ranges of p_u^0 and of $u(p^0)$, in the interval $0 \leq c \leq 0.5$, may be examined. Suppose $k = 0.10$,

Table 8. *The effects of linkage and of the value of k on the optimum mutation rate p_u^0 and its corresponding maximum $u(p^0)$ of the function $u(p;\, c, m, n)$. Here $m = 2$, $n = 3$*

c	$k = 0.01$		$k = 0.05$		$k = 0.10$		$k = 0.25$		$k = 0.30$		$k = 0.50$	
	p_u^0	$u(p^0)$	p_u^0	$u(p^0)$	p_u^0	$u(p^0)$	p_u^0	$u(p^0)$	p_u^0	$u(p^0)$	p_u^0	$u(p^0)$
0.00	.335	.0000049	.341	.000118	.349	.00045	.371	.0024	.378	.0034	.403	.0080
0.10	.441	.0000080	.449	.000193	.460	.00073	.489	.0040	.498	.0055	.531	.0131
0.20	.530	.0000113	.540	.000272	.553	.00104	.588	.0057	.599	.0078	.639	.0185
0.25	.566	.0000130	.577	.000312	.590	.00119	.628	.0065	.640	.0089	.683	.0211
0.30	.596	.0000145	.608	.000349	.622	.00133	.661	.0073	.674	.0100	.719	.0237
0.40	.641	.0000173	.654	.000416	.669	.00159	.711	.0087	.724	.0119	.772	.0282
0.50	.670	.0000196	.683	.000470	.699	.00179	.742	.0098	.756	.0135	.805	.0318

then when $n = 2$, from table 5, p_u^0 lies in the interval $0.52 \leq p_u^0 \leq 1.00$ and the maximum of u in $0.00057 \leq u(p^0) \leq 0.00226$, while for $n = 3$, from table 8, we have $0.35 \leq p_u^0 \leq 0.70$ and $0.00045 \leq u(p^0) \leq 0.00179$. The widths of the ranges of p_u^0 and of $u(p^0)$ decrease as n goes from 2 to 3; and this is true for all k. However, this does not mean that the effect of linkage on p_u^0 and $u(p^0)$ diminishes as n becomes larger. The differences between linkage and independent recombination reflected in p_u^0 and $u(p^0)$ remain important when n increases. This will be shown in Part II of this series.

$\omega(p; c, m_1, 0, m, 3)$

The special cases of this family, $m_1 = m$, constitute the set of functions $v(p; c, m, 3)$.

There is a close similarity between the functions $u(p; c, 2, 3)$ and $v(p; c, 2, 3)$ and the remarks made above for the u-function with respect to c may be extended to the function $v(p; c, m, 3)$.

As for n, its effects on p_v^0 and $v(p^0)$ are similar to those on p_u^0 and on $u(p^0)$ discussed above.

Further, we note that equation (42) depends on k whereas (41) does not. The significance of this is the same for the case where $n = 2$, described earlier.

Finally, the maximum $v(p^0; c, m, 3)$ decreases exponentially as m increases while the optimum $p_v^0(c, 3)$ is independent of m (see expressions (42) and (41)).

Acknowledgements

The author wishes to express his appreciation to Dr. Ralph E. Comstock for his stimulating suggestions during this work. He also wishes to thank Dr. James F. Crow for his advice and support during the preparation of this paper.

Literature

1. Brock, R. D.: Induced mutations affecting quantitative characters. Radiation Botany **5** (supplement), 451–464 (1965). — 2. Cooper, W. E., Gregory, W. C.: Radiation-induced leaf spot resistant mutants in the peanut. Agron. Journ. **52**, 1–4 (1960). — 3. Diamantis, B.: An application of the generating function to the solution of a genetic problem. J. Appl. Probability (1973 in press). — 4. Frey, K. J.: Mutation breeding for quantitative attributes. Radiation Botany **5** (supplement), 465–475 (1965). — 5. Gaul, H.: Use of induced mutations in seed-propagated species. Mutation and plant breeding. Nat. Acad. Sci. Publication **891**, 206–251 (1961). — 6. Gaul, H.: The concept of macro- and micro-mutations and results on induced micro-mutations in barley. Radiation Botany **5** (supplement), 407–428 (1965). — 7. Goodspeed, T. H., Olson, A. R.: The production of variation in *Nicotiana* species by X-ray treatment of sex-cells. Proc. Nat. Acad. Sci. U.S. **14**, 66–69 (1928). — 8. Gregory, W. C.: X-ray breeding of peanuts. Agron. Journ. **47**, 396–399 (1955). — 9. Gregory, W. C.: Induction of useful mutations in the peanut. Genetics in Plant Breeding, Brookhaven Symposia in Biology **9**, 177–190 (1956). — 10. Gustafsson, A.: Mutations in agricultural plants. Hereditas **33**, 1–100 (1947). — 11. Gustafsson, A.: Productive mutations induced in barley by ionizing radiations and chemical mutagens. Hereditas **50**, 211–263 (1963). — 12. Gustafsson, A., Hagberg, A., Lundquist, U.: The induction of early mutants in Bonus barley. Hereditas **46**, 675–699 (1960). — 13. Muller, H. J.: Artificial transmutation of the gene. Science **66**, 84–87 (1927). — 14. Muller, H. J.: The production of mutations by X-rays. Proc. Nat. Acad. Sci. U.S. **14**, 714–726 (1928). — 15. Muller, H. J., Herskowitz, I. H., Abrahamson, S., Oster, I. I.: A non-linear relation between X-ray dose and recovered lethal mutations in *Drosophila*. Genetics **39**, 741–749 (1954). — 16. Newcombe, H. B.: The timing of induced mutations in streptomyces. Mutation BNL **350**, 88–102 (1955). — 17. Pfeifer, R. P.: The use of an induced mutation to develop a winter barley variety. Radiation Botany **5** (supplement), 573–578 (1965). — 18. Scossiroli, R. E.: Value of induced mutations for quantitative characters in plant breeding. Radiation Botany **5** (supplement), 442–450 (1965). — 19. Stadler, L. J.: Mutation in barley induced by X-rays and radium. Science **68**, 186–187 (1928). — 20. Stadler, L. J.: Chromosome number and the mutation rate in *Avena* and *Triticum*. Proc. Nat. Acad. Sci. U.S. **15**, 876–881 (1929). — 21. Stadler, L. J.: Some genetic effects of X-rays in plants. J. Heredity **21**, 3–19 (1930). — 22. Tavcar, A.: Gamma-ray irradiation of seeds of wheat, barley and inbreds of maize and the formation of some useful point mutations. Radiation Botany **5** (supplement), 159–174 (1965).

Received January 3, 1973
Communicated by R. Riley

Dr. Basil Diamantis
Department of Medical Genetics
The University of Wisconsin
Madison, Wisconsin 53706 (USA)