Arunachalam, V.
Two Interesting Cases of Continued Mating in Monogenic and Digenic Systems.
Current Science, 68 (1993): 86-89.

Reprinted from *Current Science*

Two interesting cases of continued mating in monogenic and digenic systems

V. Arunachalam

Division of Genetics, Indian Agricultural Research Institute,
New Delhi 110 012, India

Mating with a heterozygote, continued across generations, has been investigated anew. An interesting single-gene system under this continued mating is presented and shown to extend, in general, to two genes. Based on the results, a possible modified biparental mating system of high practical significance is highlighted. Likewise, under continued selfing, the mean value of a quantitative trait has been derived, for the first time, in terms of recombination fraction and additive–additive interaction in a two-gene system.

UNDERSTANDING the dynamic flow of gene and genotypic frequencies under specific systems of mating continued across generations is very essential in predicting and directing population improvement. The stochastic process of gene frequency changes is easy to work out in terms of recurrence relationships. Such changes under popular systems of mating like selfing and sibbing continue to be the topics of discussion in many text books[1–4]. For instance, Kempthorne[2] has discussed the case of full-sib mating with any number of alleles at a single locus. He identified seven general types of mating and worked out the changes in genotypic frequencies which turned out to be complex, though elegant. Yet it would be useful for practical biologists if the theoretical logic is underlined using simple models like that of a single diallelic locus. Having built such a basic foundation, it is equally important to illustrate the complexities that would arise in cases of two or more genes; this would arrest the temptation to generalize results from one gene to many genes oblivious of the hurdles in the process. This paper aims to highlight two interesting case studies – one in single and the other in two genes – that have not been dealt with in published literature to the best of the author's knowledge. Incidental information of value to practical plant/animal breeding is also sought to be retrieved from these studies.

Case 1. Continued mating to heterozygote

In the process of illustrating Baye's theorem of 'inverse' probability, Smith[5] has described a system of mating that lends itself to illuminating an interesting phenomenon.

Initially, a cross is made between two heterozygotes. The resulting progeny is mated to a heterozygote and this process is continued. We study the population structure in successive generations.

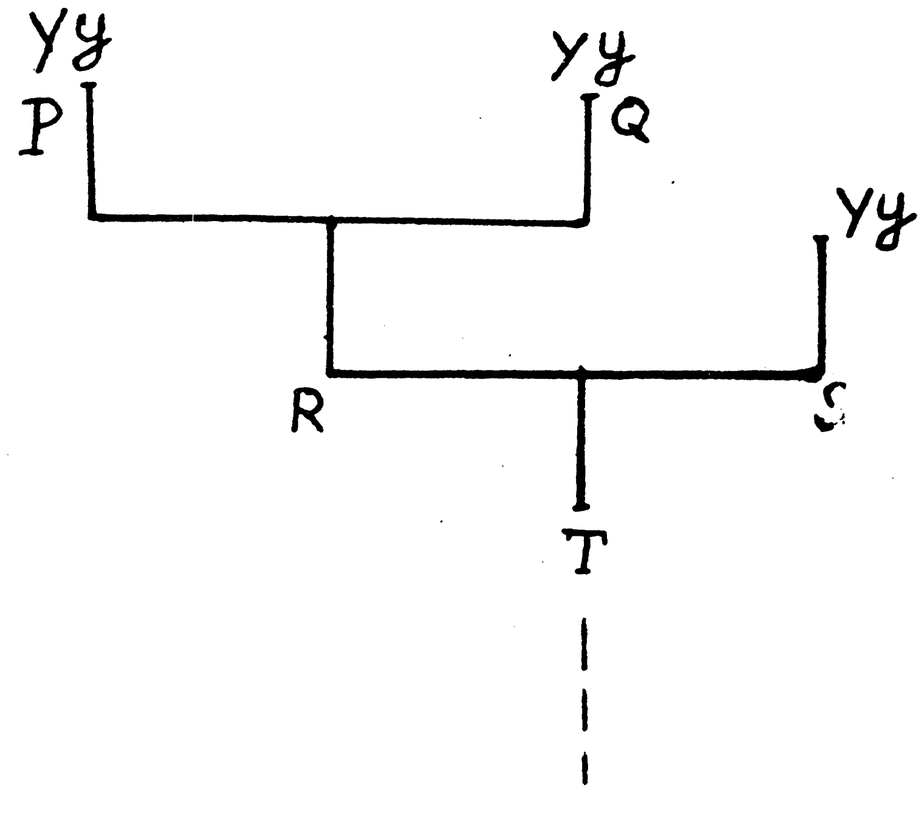

Figure 1. Continued mating to a heterozygote.

The process is illustrated in Figure 1. Let A_i $(i = 1, 3)$ denote, respectively the events that R is YY, Yy and yy. Likewise, let B_i $(i = 1, 3)$ denote, respectively the events that T is YY, Yy and yy.

Denote $P_r(R = A_i) = x_i$

$$P_r(T = B_i) = Z_i \quad (i = 1, 3)$$

$$P_r(T = B_r \mid R = A_s) = m_{rs} \quad (r, s = 1, 3)$$

The matrix (m_{rs}) is obviously given by

$$\boldsymbol{M} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

From elementary principles, it is easily seen, for instance, that

$$Z_2 = P_r(T = B_2) = x_1 m_{21} + x_2 m_{22} + x_3 m_{23}$$

$$= \frac{1}{4}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{4}\cdot\frac{1}{2} = \frac{1}{2}$$

Thus, if $\boldsymbol{Z}_1' = (z_1, z_2, z_3)$

and

$$\boldsymbol{X}' = (x_1, x_2, x_3)$$

then $\boldsymbol{Z}_1 = \boldsymbol{MX}$

$$= \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}$$

This state is identical to the initial state. Therefore, the frequencies of possible genotypes in the second generation will again be given by $\boldsymbol{Z}_2 = \boldsymbol{MZ}_1 = \boldsymbol{M}\,\boldsymbol{MX} = \boldsymbol{M}^2\boldsymbol{X}$, where

$$\boldsymbol{M}^2 = \begin{bmatrix} \frac{3}{8} & \frac{1}{4} & \frac{1}{8} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{8} & \frac{1}{4} & \frac{3}{8} \end{bmatrix}$$

Likewise, $\boldsymbol{Z}_3 = \boldsymbol{M}^3\boldsymbol{X}$, where

$$\boldsymbol{M}^3 = \begin{bmatrix} \frac{5}{16} & \frac{1}{4} & \frac{3}{16} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{16} & \frac{1}{4} & \frac{5}{16} \end{bmatrix}$$

Note that the third row is the first row in reverse order and the second row is constant in every generation.

The element m_{11} of matrix $\boldsymbol{M}$ in successive generations is given by 2/4, 3/8, 5/16, 9/32, 17/64, ... in other words,

t_n = value of m_{11} (or m_{33}) in generation n

$$= \frac{1}{4} + \frac{1}{2^{n+1}} \quad \text{for} \quad n = 1, 2, \ldots$$

Similarly, U_n = value of m_{13} (or m_{31}) in generation n

$$= \frac{1}{2} - t_n$$

Thus, $\boldsymbol{Z}_n = \boldsymbol{M}^n\boldsymbol{X}$.

As $n \to \infty$, $t_n \to 1/4$ and $U_n \to 1/2$, the giving

$$\boldsymbol{M}^\infty = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

leading to $\boldsymbol{Z}_\infty = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{bmatrix}$

In general, Hardy–Weinberg equilibrium is maintained under six possible matings (random mating). But in this case, restricted mating to a heterozygote alone, viz. YY × Yy : Yy × Yy : yy × Yy, maintains such an equilibrium with initial frequencies of genotypes (1/4, 1/2, 1/4).

To examine whether the results extend to a general two-gene case with linkage, let us consider an F_1 generation consisting of only AB/ab, the coupling double heterozygote. Let the recombination fraction be r. As described earlier, let F_2 be obtained by mating F_1 with AB/ab and F_3 by mating the progeny generation with AB/ab and so on (Table 1). On examining the frequency of the 10 resulting genotypes, it is easily seen that they can be arranged in 5 genotypic groups. Individual genotypes in each group have identical frequencies (see Table 1). Further, we note that

Table 1. Frequency of genotypes under heterozygote × heterozygote matings in two genes.

$$\frac{AB}{ab}F_1 \times \frac{AB}{ab}$$

$$\downarrow$$

$$F_2 \times \frac{AB}{ab}$$

leads to

five groups of genotypes

I. $\frac{AB}{AB}, \frac{ab}{ab}$

II. $\frac{AB}{Ab}, \frac{AB}{aB}, \frac{Ab}{ab}, \frac{aB}{ab}$

III. $\frac{Ab}{Ab}, \frac{aB}{aB}$

IV. $\frac{AB}{ab}$

V. $\frac{Ab}{aB}$

Genotypic group	Frequency in F_2	Frequency in F_3
I	$\frac{1}{4}(1-r)^2$	$\frac{1}{8}(1-r)[(1-r)(1-2r)+1]$
II	$\frac{1}{2}r(1-r)$	$\frac{1}{2}r[(1-r)^2+1/4]$
III	$\frac{1}{4}r^2$	$\frac{1}{8}r^2(3-2r)$
IV	$\frac{1}{2}(1-r)^2$	$\frac{1}{4}(1-r)[(1-r)(1-2r)+1]$
V	$\frac{1}{2}r^2$	$\frac{1}{4}r^2(3-2r)$

Table 2. Transition of genotypic group frequencies under linkage in two genes.

Genotypic group*	$r = 0$		$r = 1/4$		$r = 1/2$	
	F_2	F_3	F_2	F_3	F_2	F_3
I	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{9}{64}$	$\frac{33}{256}$	$\frac{1}{16}$	$\frac{1}{16}$
II	0	0	$\frac{3}{32}$	$\frac{13}{128}$	$\frac{1}{8}$	$\frac{1}{8}$
III	0	0	$\frac{1}{64}$	$\frac{5}{256}$	$\frac{1}{16}$	$\frac{1}{16}$

*As in Table 1.

genotypic groups IV and V consist of coupling and repulsion double heterozygotes with twice the frequencies of groups I and III, respectively. Hence, for all practical purposes, it is enough to consider groups I, II and III only, whose frequencies in F_2 and F_3 have been derived (Table 1).

Three distinct cases of r ($r = 0, 1/4, 1/2$) help to evaluate, under linkage, the trend in transition of frequencies from F_2 to F_3 in various groups (Table 2). When linkage is extremely strong ($r \simeq 0$), the population consists of only AB/AB, AB/ab and ab/ab which is identical to the one-gene case, gametes AB, ab behaving like alleles of a single locus. In the case of $r = 1/2$, equivalent to no linkage, the population remains in Hardy–Weinberg equilibrium. Therefore, this case is also analogous to the monogenic case. When the linkage is not tight ($0 < r < 1/2$), the frequencies of groups II and III, and, therefore, of V increase at the cost of I and IV. In other words, frequencies of homozygotes at both loci (AB/AB, ab/ab) and of the coupling double heterozygote (AB/ab) decrease, consequently increasing heterozygosis. This result is similar to that derived in the single-gene case. Thus, the results of the monogenic case hold broadly in the digenic case too. It is easy to visualize the corresponding results if we start with a repulsion double heterozygote in F_1.

The mating system described above can be practised in F_2 and further generations in breeding for improved productivity in plants. At present even biparental mating (in F_2) that neither distinguishes parental genotypes nor restricts matings to heterozygous parents has been established to be superior than selfing in producing utilizable genetic variability. In this context, the present results focus sure clues to improve such variability still further. Even in polygenic systems (more relevant to practical plant breeding situations), it is possible to detect more likely heterozygotes. For instance, such genotypes would show significant improvement in the desired direction over the better parent for a number of characters (as per heterosis concepts). Restricting the mating of the progeny population to such genotypes would scale up the frequency of matings involving a heterozygote. Further, by similar methods one can restrict the parents chosen from the progeny population to the most-likely heterozygotes. Only more experiments with plant/animals can provide confirmatory support and crystallize concepts with concrete application potential.

Table 3. QT means in successive generations of selfing in two linked genes.

Generation	QT mean
F_2	$\frac{1}{2}(h_a + h_b) + ix + \frac{1}{2}ly$
F_3	$\left(z + \frac{1}{4}y^2\right)(h_a + h_b) + iw(1 - x^2) + \frac{1}{4}ly^2$
...	
F_n	$\left[\left(\frac{1}{2}\right)^{n-3} Z + \left(\frac{1}{2}\right)^{n-1} y^2\right] \times (h_a + h_b) + iw(1 - x^{n-1}) + \left(\frac{1}{2}\right)^{n-1} ly^{n-1}$
F_∞	iw

$x = 1/2 - r$; $w = (1/2 - r)/(1/2 + r)$; $y = r^2 + (1 - r)^2$;
$z = r(1 - r)(1 - r + r^2)$.

Case 2. Continued selfing in two linked genes

Selfing is the most common method of advancing generations in plant breeding. The reduction of heterozygosity under selfing is well-known and recurrence relations have been developed for a single gene and two unlinked genes in the absence of linkage disequilibrium. In the latter case, the process is equivalent to a single gene except that there are two dominance effects (h_a, h_b) corresponding to the two genes and consequently a dominance–dominance interaction (l). The theory has been well documented[2, 4].

It is generally believed that the final state, namely, complete absence of heterozygosity, will be reached asymptotically when factors like linkage disequilibrium, inbreeding and low recombination (recombination fraction r near zero) implying tight linkage are taken into account. The mean value for a quantitative trait (QT) decreases gradually over successive generations. In a single gene, QT mean in the final state becomes equal to $(p - q)d$, where p is the frequency of the dominant allele, $q = 1 - p$ and d, h, $-d$ represent the QT values of dominant homozygote, heterozygote and recessive homozygote, respectively, following Mather and Jinks[4]. If, in addition, $p = 1/2$, the QT mean at the final state = 0.

In the case of two genes, this result does not extend unconditionally. When the recombination fraction $r = 1/2$ and initial linkage disequilibrium D is absent in large random mating populations, an equilibrium is maintained. On the other hand, continuous selfing will lead to a final state in which the four homozygotes, viz.

Table 4. Changes in mean values of QT over successive generations of selfing in two linked genes for various values of recombination.

Generation	$r=\frac{1}{2}$	$r=\frac{1}{4}$	$r=0$
F_1	h_a+h_b+l	h_a+h_b+l	h_a+h_b+l
F_2	$\frac{1}{2}(h_a+h_b)+\frac{1}{4}l$	$\frac{1}{2}(h_a+h_b)+\frac{1}{2}\left(\frac{5}{8}\right)l+\frac{1}{4}i$	$\frac{1}{2}(h_a+h_b)+\frac{1}{2}l+\frac{1}{2}i$
F_3	$\frac{1}{4}(h_a+h_b)+\frac{1}{16}l$	$\frac{1}{4}(h_a+h_b)+\frac{1}{4}\left(\frac{5}{8}\right)^2 l+\frac{5}{16}i$	$\frac{1}{4}(h_a+h_b)+\frac{1}{4}l+\frac{3}{4}i$
F_n	$\frac{h_a+h_b}{2^{n-1}}+\frac{l}{4^{n-1}}$	$\frac{(h_a+h_b)+(5/8)^{n-1}l}{2^{n-1}}+\frac{1}{3}\left(1-\frac{1}{4^{n-1}}\right)i$	$\frac{h_a+h_b+l}{2^{n-1}}+\left(1-\frac{1}{2^{n-1}}\right)i$
F_∞	0	$\frac{1}{3}i$	i

AABB, AAbb, aaBB and aabb (if the two genes are A–a, B–b) will occur with equal frequencies if $p_1 = p_2 = 1/2$, leading to a QT mean value = 0. This result is then an extension of the result from a single gene.

However, when the recombination fraction ≠ 1/2, such results do not extend. The recurrence relations giving the QT mean value in successive generations, not published so far, have now been derived and presented in Table 3. The QT mean is shown to have contributions of not only h_a, h_b and l but also i, the additive–additive interaction effect.

When $r = 0$, i.e. when there is complete linkage, the QT mean in the final state (after a very large number of generations corresponding to F_∞) = i, which goes on reducing and becomes zero when $r = 1/2$ (corresponding to unlinked loci, as mentioned earlier). Broad extrapolation of these results to a general polygenic case would suggest that the QT mean, after a very large number of generations, would still be a function of interactions, including the additive–additive type, for the two-gene case (Table 4). In practice, it is not possible to attain the final state F_∞. In addition, plant breeders advance, most often, progeny generations by a finite sample of selfings, which in reality, may lead to results far away from the theoretical expectations (cf. Tables 3 and 4).

It is thus clear that practical plant breeding situations suffer a few drawbacks:

(a) The number of matings to advance progeny generations is finite and often too small to be amenable to theoretical formulation. Further, the finite and small progeny sizes maintained in practical breeding experiments cannot admit to a large number of successive selfings and, therefore, the F_∞ state can rarely be approached.

(b) Assumptions of a large population, independent and additive genes on which theories are built[4] cannot, in general, fit practical plant breeding situations, grossly reducing the suitability of such theories.

Nevertheless, the result that in a two-gene system practising selfing the QT mean at the final state would be a function of the value of additive–additive effect and r, the recombination fraction, is significant.

1. Li, C. C., *Population Genetics*, Chicago University Press, Chicago, 1955.
2. Kempthorne, O., *An Introduction to Genetic Statistics*, John Wiley & Sons, New York, 1957.
3. Crow, J. F. and Kimura, M., An Introduction to Population Genetics Theory, Harper and Row, New York, 1970.
4. Mather, K. and Jinks, J. L., *Biometrical Genetics*, Chapman and Hall, London, 1971.
5. Smith, C. A. B., *Biomathematics, Vol. II*, Charles Griffin & Co., London, 1969.

Received 2 November 1993; revised accepted 10 October 1994