Journal of Genetics Volume 63, Number 1 April 1977

HETEROSIS FOR CHARACTERS GOVERNED BY TWO GENES

V. Arunachalam
Division of Genetics, Indian Agricultural Research Institute
New Delhi, India

SUMMARY

An expression for heterosis over midparent is derived for characters governed by one gene when hybrids are made using arbitrarily inbred parents. It is shown that, so long as inbreeding does not depress the vigour and fitness of parents, heterosis in hybrids between them will be greater than that obtained in hybrids between open-pollinated parents.

The expression derived for heterosis for characters governed by two genes has brought to focus the possibilities of obtaining heterosis without the presence of dominance effects, and also of dominance interactions. Arguments are advanced to illustrate that pure dominance hypothesis alone cannot account for the observed heterosis when the character is under the control of more than one gene. It is found feasible that pure additive gene action at individual genic level coupled with a favourable additive x additive interaction can produce heterosis when gene frequency differences can be ensured. Implications of this result to problems in plant breeding are discussed.

INTRODUCTION

Heterosis, expressed as the improvement in a character shown by a hybrid over their mid- or better parental value, is a vital measure of the genetic progress made in plant and animal selection. Several hypotheses have been advanced to explain this phenomenon (Falconer, 1964; Li, 1955; Sinha and Khanna, 1975). In particular, the analytical treatment of heterosis in hybrids between populations governed by a single diallelic gene given by Falconer deserves special mention. He has shown that heterosis is equal to the dominance effect multiplied by the square of the difference in the dominant allelic frequency. This hypothesis has formed the basis for plant breeders to use genetically diverse parents in the production of heterotic hybrids. Such a procedure will result in heterosis if genetic divergence is associated, as is usually found to be, with a gene frequency difference. However, the amount of heterosis that is obtained cannot be predicted by the degree of genetic divergence between the parents, as many studies in crop plants have shown (see, for example, Timothy, 1963 in maize). These concepts have been carefully studied by Cress (1966) in a multi-allelic single gene system.

During an evaluation of single and multiple cross hybrids in triticale and pearl

millet, it was observed that heterosis can result by the operation of general combining ability effects (g. c. a.) alone in the desired direction. As Falconer (1964) has shown, g. c. a. of a parent (in a single gene case) is a relative measure of its additive (or breeding) value. This has raised the following interesting questions : (a) What is the relative importance of gene frequency difference and dominance in determining heterosis under single gene control? (b) Can heterosis result solely due to the consonant interaction of additive effects of genes when more than one gene is considered? An attempt is made in this paper to answer them. Throughout our discussion hereafter, we shall consider the expression, H = (Hybrid mean - Midparental value), to discuss the nature of heterosis, though the actual expression for heterosis (%) would be H_O = (C - P) x 100/P where C = Hybrid mean, P = Midparental value. However, this minor change will not affect the validity of the results related to heterosis.

THEORY

(i) Single gene diallelic population : Let us consider two populations I and II, each containing the genotypes, AA, Aa and aa. Let, in population I, the gene frequency of A = p and in II, the gene frequency of A = p' = p - y so that y is the difference between the gene frequencies of A in the two populations. Let F and F' be the inbreeding coefficients of the populations I and II so that the initial frequencies of the three genotypes are :

Population	AA	Aa	aa
I	$p^2 + Fpq$	$2pq\ (1 - F)$	$q^2 + Fpq$
II	$p'^2 + F'p'q'$	$2p'q'\ (1 - F')$	$q'^2 + F'p'q'$

Following the symbolic algebra approach (Arunachalam and Owen 1971), let us represent the values of quantitative character of the genotypes by the genotypic symbols themselves. Then, the population mean of I, MP1 = $(Ap + aq)^2$ + $Fpq\ (A - a)^2$, which is a symbolic representation, the genotypic symbols attain meaning only when expanded and when the actual values are substituted. Similarly, the population mean of II, MP2, will be given by

$$MP2 = (Ap' + aq')^2 + F'p'q'\ (A - a)^2$$

$$= (Ap + aq)^2 + [(F + F')/2]pq\ (A - a)^2 - y\ (A - a)(Ap + aq) +$$

$$(y^2/2)\ (A - a)^2\ (1 - F') + (y/2)F'\ (p - q)\ (A - a)^2$$

Noting that L = additive effect = $(A - a)\ (Ap + aq)$ and
Q = dominance effect = $(A - a)^2$,

if, using Mather's notation, we substitute d, h, and -d for the values of the

quantitative character for AA, Aa and aa respectively, it is easy to see that $L = d + h(q - p)$ and $Q = -2h$. Hence, the midparental value = $M_0 - y[d + h(q - p)] - h[pq(F + F') + yF'(p - q) + y^2(1 - F')]$ where M_0 is the population mean under complete random mating. Under random mating of populations I and II, the F1 population would have genotypes AA, Aa and aa with respective frequencies $p(p - y)$, $2pq + y(p - q)$ and $q(q + y)$ (since inbreeding does not change the gene frequencies). The F1 population mean,

$$MF1 = (Ap + aq)^2 - y(A - a)(Ap + aq) = M_0 - y[d + h(q - p)]$$

$$H = -(Q/2)[(F + F')pq + yF'(p - q) + y^2(1 - F')]$$

$$= h[(F + F')pq + yF'(p - q) + y^2(1 - F')]$$

When $F = F' = 0$, i.e., under complete random mating, heterosis = hy^2, as shown by Falconer (1964). Some particular cases are of special interest. When $F = F'$, heterosis = $h[2Fpq + yF(p - q) + y^2(1 - F)]$ which can easily be shown to be $> hy^2$, for positive dominance effect and non-null values of p and p'. If $F = F' = 1$, i.e., for completely inbred populations, heterosis = $h[2pq + y(p - q)]$ and = $h/2$ if $p = q = 1/2$, in addition, as has been shown by Mather and Jinks (1971) by considering the cross, AA x aa.

(ii) Two gene populations : Let the gametic frequencies in the two populations governed by the genes A_1, a_1 ; A_2, a_2 be as follows :

Population	A_1A_2	Gametic frequencies of a_1a_2	A_1a_2	a_1A_2
I	P_1	P_2	P_3	P_4
II	$P_1 + x$	$P_2 + y$	$P_3 + z$	$P_4 - x - y - z$

The frequencies P_1, P'_1 of the genes A_1, A_2 and of their alleles will then be :

$P_1 = P_1 + P_3$ $\quad P'_1 = P_1 + x + z$ $\quad q'_1 = q_1 - x - z$

$P_2 = P_2 + P_4$ $\quad P'_2 = P_2 - y - z$ $\quad q'_2 = q_2 + y + z$

Let D and D' be the amounts of linkage disequilibrium in the populations I and II.

$$D' = (P_1 + x)(P_2 + y) - (P_3 + z)(P_4 - x - y - z)$$

$$= D + xq_2 + yp_1 + z(p_1q_2 - p_2q_1) + (p'_1 - p_1)(q'_2 - q_2), \text{ after some algebra.}$$

Let $p'_1 - p_1 = g_1$; $p'_2 - p_2 = g_2$; $q'_1 - q_1 = f_1$; $q'_2 - q_2 = f_2$

$x + z = g_1 = -f_1$ and $y + z = -g_2 = f_2$ so that $g_1 + f_1 = g_2 + f_2 = 0$

It can be easily shown that

$$xq_2 + yp_1 + z(p_1q_2 - p_2q_1) = g_1q_2 + f_2p_1 - z \text{ and}$$

$$D' = D + g_1q_2 - g_2p_1 - z - g_1g_2$$

As before, let the quantitative character be represented by the genotypic symbols; we use the notation of Arunachalam and Owen (1971) and the symbolic algebra approach to derive the expression for heterosis.

$$S_1 = A_1p_1 + a_1q_1$$

$$S_2 = A_2p_2 + a_2q_2$$

$$U = S_1S_2 + D\,(A_1 - a_1)\,(A_2 - a_2)$$

$$S'_1 = A_1\,(p_1 + x + z) + a_1\,(q_1 - x - z)$$

$$= S_1 + g_1\,(A_1 - a_1)$$

Similarly, $S'_2 = S_2 + g_2\,(A_2 - a_2)$

$$U' = [S_1 + g_1(A_1 - a_1)][S_2 + g_2\,(A_2 - a_2)] + (D + g_1q_2 - g_2p_1 - z - g_1g_2)(A_1 - a_1)(A_2 - a_2)$$

Using the relations, $A_2 = S_2 + q_2\,(A_2 - a_2)$ and $a_1 = S_1 - p_1\,(A_1 - a_1)$,

we get $U' = U + g_1A_2\,(A_1 - a_1) + g_2a_1(A_2 - a_2) - z\,(A_1 - a_1)(A_2 - a_2)$ .... (E1)

We now represent the genotypic values of the quantitative character of the individuals by w_{ij} (i, j = 1, 4) as shown below :

	A_1A_2	a_1a_2	A_1a_2	a_1A_2
A_1A_2	w_{11}	w_{12}	w_{13}	w_{14}
a_1a_2	w_{12}	w_{22}	w_{23}	w_{24}
A_1a_2	w_{13}	w_{23}	w_{33}	w_{34}
a_1A_2	w_{14}	w_{24}	w_{34}	w_{44}

Where $w_{12} = w_{34}$

The mean value of the population I = v = U^2

The mean value of the population II = V' = U'^2

$$= U^2 + g_1^2(w_{11} - 2w_{14} + w_{44}) + 2g_1g_2(w_{14} + w_{24} - w_{34} - w_{44}) + g_2^2(w_{22} - 2w_{24} + w_{44})$$
$$+ 2g_1(w_{1.} - w_{4.}) + 2g_2(w_{4.} - w_{2.}) - 2z(L_{12} + DQ_{12}) - 2g_1z(w_{11} + 2w_{12} - w_{13} -$$
$$2w_{14} - w_{24} + w_{44}) - 2g_2z(w_{14} + 2w_{24} - w_{44} - w_{22} + w_{23} - 2w_{12}) + z^2Q_{12}$$

where L_{12}, Q_{12} are the additive x additive and dominance x dominance interaction effects defined by Arunachalam and Owen (1971).

If $\epsilon_i = w_{i1} + w_{i2} - w_{i3} - w_{i4}$ (i = i, 4) are the epistatic components,

$$V' = U^2 + g_1^2(w_{11} - 2w_{14} + w_{44}) + 2g_1g_2\epsilon_4 + g_2^2(w_{22} - 2w_{24} + w_{44}) + 2g_1(w_{1.} - w_{4.}) +$$
$$2g_2(w_{4.} - w_{2.}) - 2z(L_{12} + DQ_{12}) - 2g_1z(\epsilon_1 - \epsilon_4) - 2g_2Z(\epsilon_4 - \epsilon_2) + z^2Q_{12}$$

where $w_{1.} = \sum_j w_{ij}$ are the marginal means,

Considering the random mating process as the random union of gametes, if the genotypic values are the same in the two populations, the mean value in F_1 generation will be

$$MF1 = UU'$$

The logical derivation of the above expression is simple once we consider the corresponding situation in the single gene case. We recall that the genotypic values of a quantitative character have been considered to be the same in the two populations, p_1, p_2 being the respective gene frequencies of A and V_1, V_2 the population mean values. Then

$$V_1 = (Ap_1 + aq_1)^2$$

$$V_2 = (Ap_2 + aq_2)^2$$

and $MF1 = (Ap_1 + aq_1)(Ap_2 + aq_2)$ which can easily be verified. This logic will extend to two gene case, U^2, U'^2 taking the places of V_1 and V_2.

Hence in the two gene case, we have

$$H = MF1 - (V + V')/2$$

$$= UU' - (U^2 + U'^2)/2$$

$$= -1/2\,(U' - U)^2$$

$$= -1/2\,[g_1A_2(A_1 - a_1) + g_2a_1(A_2 - a_2) - z(A_1 - a_1)(A_2 - a_2)]^2 \quad \text{by (E1)}$$

Expanding the expression and simplifying, we get

$$H = -1/2 \quad [g_1^2(w_{11} - 2w_{14} + w_{44}) + g_2^2(w_{44} - 2w_{24} + w_{22}) + z^2 Q_{12} + 2g_1 g_2 (w_{14} - w_{12} - w_{44} + w_{24})$$
$$- 2g_1 z(w_{11} - 2w_{14} + w_{44} - w_{13} + 2w_{12} - w_{24}) - 2g_2 z(w_{14} - 2w_{12} + w_{23} - w_{44} - w_{22} + 2w_{24})]$$

A study of the expressions of various genetic effects given in Arunachalam and Owen (1971, p. 12), shows that the dominance effect at locus 1 can be written as

$$Q_1 = p_2^2 Q_{A_2A_2} + 2p_2q_2 Q_{A_2a_2} + q_2^2 Q_{a_2a_2}$$

where $Q_{A_2A_2}$, for example, is the dominance effect corresponding to the genotype A_2A_2 corresponding to gene A_2. Writing the character values in the usual 3 x 3 matrix from corresponding to the three genotypes at each locus, it is easy to see that

$$Q_{A_2A_2} = 1/2 \quad \frac{\partial^2}{\partial p_1^2} (w_{11}p_1^2 + 2w_{14}p_1q_1 + w_{44}q_1^2) = w_{11} - 2w_{14} + w_{44}$$

Thus,

$$H = -1/2 \; [g_1^2 Q_{A_2A_2} + g_2^2 Q_{a_1a_1} + z^2 Q_{12} + 2g_1g_2 \epsilon_{a_1A_2} - 2g_1 z(Q_{A_2A_2} - Q_{A_2a_2})$$
$$- 2g_2 z(Q_{A_1a_1} - Q_{a_1a_1})]$$

where

$$\epsilon_{a_1A_2} = \epsilon_4 \text{ defined earlier} \qquad \ldots . (E2)$$

Following Mather and Jinks (1971), the character values can be written in terms of d_k, h_k (k = 1, 2), i, j_{12}, j_{21} and ℓ where the genotypes A_kA_k, A_ka_k and a_ka_k have the values d_k, h_k and $-d_k$ (k = 1, 2), and i = homozygote x homozygote, j_{12} = homozygote x heterozygote, j_{21} = heterozygote x homozygote and ℓ = heterozygote x heterozygote interaction effects.

Then, on simplification,

$$H = g_1^2(h_1 + j_{21}) + g_2^2(h_2 - j_{12}) - 2\ell z^2 + g_1g_2(\ell - i - j_{21} + j_{12}) + 2g_1 z(\ell - j_{21}) - 2g_2 z(\ell + j_{12}) \qquad (E3)$$

It is easily shown that i, j_{12}, j_{21} and ℓ do not represent the interaction effects but a function of them do. In fact, substitution of the values of the character in the expressions given in Arunachalam and Owen (1971) provide the values shown in Table 1.

Table 1. Expressions for the various genetic effects and interactions

EFFECT	EXPRESSION	
	a	b
Additive (locus 1) L_1	$d_1+r_1h_1-r_2i-r_1r_2j_{21}+2s_2j_{12}+r_1s_2\ell-2D(r_2\ell+j_{21})$	$d_1+\frac{1}{2}j_{12}$
Additive (locus 2) L_2	$d_2+r_2h_2-r_1i-r_1r_2j_{12}+2s_1j_{21}+2r_2s_1\ell-2D(r_1\ell+j_{12})$	$d_2+\frac{1}{2}j_{21}$
Dominance (locus 1) O_1	$-2h_1+2r_2j_{21}-4s_2\ell$	$-2h_1-\ell$
Dominance (locus 2) O_2	$-2h_2+2r_1j_{12}-4s_1\ell$	$-2h_2-\ell$
Additive x Additive L_{12}	$i+r_1j_{21}+r_2j_{12}+(r_1r_2+2D)\ell$	i
Additive x Dominance L_1O_2	$-2j_{12}-2r_1\ell$	$-2j_{12}$
Dominance x Additive L_2O_1	$-2j_{21}-2r_2\ell$	$-2j_{21}$
Dominance x Dominance O_{12}	4ℓ	4ℓ

a = general ; b = when $p_1 = p_2 = 0.5$, $D = 0$; $r_1 = q_1 - p_1$; $r_2 = q_2 - p_2$;

$s_1 = p_1q_1$; $s_2 = p_2q_2$

When D = 0, it is known that the total genetic variance for the character is given by,

$$VG = 2s_1L_1^2 + 2s_2L_2^2 + s_1^2Q_1^2 + s_2^2Q_2^2 + 4s_1s_2L_{12}^2 + 2s_1s_2^2(L_1Q_2)^2 + 2s_1^2s_2(L_2Q_1)^2 + s_1^2s_2^2Q_{12}^2$$

where $s_k = P_kq_k$ (k = 1, 2) (Cockerham, 1954; Arunachalam and Owen, 1971).

If, in addition, $p_1 = p_2 = 0.5$, we get

$$VG = 1/2\ (d_1 + \frac{1}{2}j_{12})^2 + \frac{1}{2}(d_2 + \frac{1}{2}j_{21})^2 + \frac{1}{4}(h_1 + \frac{1}{2}\ell)^2 + \frac{1}{4}(h_2 + \frac{1}{2}\ell)^2 + \frac{1}{4}i^2 + \frac{1}{8}j_{12}^2 + \frac{1}{8}j_{12}^2 + \frac{1}{16}\ell^2$$

as shown by Mather and Jinks (1971, p.172). The implications of the expression for H in the two gene case have to be considered in the light of the above observations.

DISCUSSION

The genetic basis of heterosis still poses a number of unsolved problems. A welcome start was made when the role of the differences in gene frequency in addition to dominance effect was illustrated in a single gene case (Falconer, 1964). An extended analysis of a single gene multiallelic case (Cress, 1966) has shown that, while the proposition -- a heterotic hybrid would entail genetically divergent parents -- is true, its converse need not always be true. The mutual cancellation of positive and negative effects of multiple alleles can thus lead to poor hybrid performance; this would be in addition to the effects of linkage and epistasis when more than one gene is involved. The evaluation of the degree of parental divergence on the basis of the heterotic response obtained is hence invalid. His results have raised the curiosity of theoretical and applied geneticists alike, to locate other genetic possibilities to obtain heterosis. Our study has shown that, other than dominance and gene frequency difference (in a single gene case), hibridization of inbred parents enhances heterosis, for which ample experimental evidence is available in a range of crop plants. But this is subject to the vital condition that inbreeding does not depress the vigour or fitness of the parents. Usually a depression in vigour is a consequence of inbreeding in cross-pollinated crops. In practice, the results of a single gene case cannot be extended to situations involving many genes. Correlation of gene frequencies between genes at linked loci and complex epistatic interactions are usually met with in dealing with characters governed by more than one gene. Sved (1972) has shown that associative overdominance, linkage disequilibrium and the selective advantage or otherwise of chromosome segments (rather than genes) play crucial roles in the expression of heterosis. Our study on heterosis in two gene diallelic systems supplements his results and brings to focus the role of epistatic interactions in causing heterosis.

The expression for H involves terms in g_1, g_2, the gene frequency differences for the genes A_1 and A_2 and z containing, in addition, the difference in linkage disequilibrium between the parental populations. Superficially H involves terms related to dominance effects of gametes (expression E2), but a closer analysis shows that it contains terms really related to interaction effects (expression E3). It can be seen that, when the constants, i, j_{12}, j_{21} and ℓ are zero, $h = g_1^2 h_2 + g_2^2 h_2$, which is a straightforward extension of Falconer's result for a single gene case.

A large number of possibilities exists for realising a significant H value in the presence of interactions alone; of particular importance, is the role of i which is the additive x additive effect when $p_1 = p_2 = 0.5$, $D = 0$ or which is the value of L_{12} when $J_{21} = j_{12} = \ell = 0$.

It can be seen that $H = - g_1 g_2 i$ when $h_1 = h_2 = j_{21} = j_{12} = \ell = 0$. This shows that it is possible to realise heterosis on the strength of gene frequency differences g_1, g_2 acting in unison with additive x additive interaction. This suggests that, theoretically, the presence of dominance effects and domiance interactions is not necessary to realise heterosis for characters governed by more than one gene. Pure additive gene action at an individual genic level coupled with a favourable additive x additive interaction can produce heterosis when the gene frequency differences are ensured. Work on single and multiple crosses in crop plants like triticale, pearl millet and turnip rape in our unit has confirmed that (a) it is possible to obtain heterosis in single crosses when the g. c. a. of parents are alone significant (with non-significant specific combining ability) (b) a choice of such F_1's as female parents with an appropriate selection of male parents, would ensure heterosis in multiple crosses with a high probability. The obvious advantage of heterosis based on additive gene action and interaction lies in the fact that it can be sustained in the hybrid derivatives in advanced generations. This would offer a new breeding procedure for a variety derivative from multiple crosses, in self-pollinated crops.

Plant breeders in general are concerned with heterosis for yield. It is now well-known that heterosis for component characters is as vital as heterosis for yield per se and hence breeding and selection techniques should aim at heterosis at component levels and its associated transfer to yield. An elaborate review of this concept has been made using a number of examples from crop plants by Sinha and Khanna (1975). They stress that heterosis at the level of components is crucial for heterosis at the level of a character which is a multiplicative product of two or more components. However, their interpretation of the genetic basis of heterosis is based on assumptions which are debatable ; for example, dominance is inferred from F_1 data on phenotypes and multiplicative gene action of component characters is assumed without reasonable evidence. Thus their hypothesis that dominance at component levels will lead to heterosis at character levels would admit of modification and sound alternatives. The present study presents, for example, a number of possibilities for obtaining heterosis solely through epistatic interactions among the genes controlling the trait.

In conclusion, we may observe that pure dominance hypothesis alone for heterosis is not valid any longer for characters governed by more than one gene and it is time that due emphasis is laid on alternative genetic mechanisms for heterosis breeding.

REFERENCES

Arunachalam, V. and Owen, A.R.G. (1971). Polymorphisms with Linked Loci. London : Chapman and Hall

Cockerham, C.C. (1954). An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. Genetics 39, 859-882.

Cress, C.E. (1966). Heterosis of the hybrid related to gene frequency differences between two populations. Genetics 53, 269-274.

Falconer, D.S. (1964). An Introduction to Quantitative Genetics (3rd Imp.). Edinburgh and London : Oliver and Boyd

Li, C.C. (1955). Population Genetics. University of Chicago Press

Mather, K. and Jinks, J.L. (1971). Biometrical Genetics. London : Chapman and Hall

Sinha, S.K. and Khanna, R (1975). Physiological, biochemical and genetical basis of heterosis. Advances in Agronomy 27, 123-174

Sved, J.A. (1972). Heterosis at the level of the chromosome and at the level of the gene. Theoretical Population Biology 3, 491-506

Timothy, D.H. (1963). Genetic diversity, heterosis and the use of exotic stocks in maize in Colombia. Statistical Genetics and Plant Breeding, NAS-NRC 982, 581-593.