JAMES F. CROW
University of Wisconsin

Chapter 18

Dominance and Overdominance[*]

Since the first attempts to explain hybrid vigor and the deleterious effects of inbreeding in Mendelian terms, there have been two principal hypotheses. Both were advanced early, and though each has had its ups and downs in popularity, both have persisted to the present time. The first hypothesis is based on the observed correlation between dominance and beneficial effect (or recessiveness and detrimental effect). Inbreeding uncovers deleterious recessives, and typically results in deterioration.

With hybridization, some of the detrimental recessives brought into the hybrid zygote by one parent are rendered ineffective by their dominant alleles from the other, and an increase in vigor is the result. If the number of factors is large, or if there is linkage, the probability becomes exceedingly small of a single inbred line becoming homozygous for only the dominant beneficial factors. Consequently, there should be a consistent decrease in vigor with inbreeding, and recovery with hybridization. This idea has been called the *dominance* or the *dominance of linked genes* hypothesis.

The alternative theory assumes that there is something about hybridity *per se* that contributes to vigor. In Mendelian terms this means that there are loci at which the heterozygote is superior to either homozygote, and that there is increased vigor in proportion to the amount of heterozygosis. This idea has been called *stimulation of heterozygosis*, *super-dominance*, *over-dominance*, *single gene heterosis*, *cumulative action of divergent alleles*, and simply *heterosis*.

In accordance with the title of this discussion I shall use the words *dominance* and *overdominance* for the two hypotheses. This leaves the word *heterosis* free for more general use as a synonym for *hybrid vigor* (Shull, 1948).

* Paper No. 434 from the Department of Genetics, University of Wisconsin.

In most situations, the hypotheses of dominance and overdominance lead to the same expectations. In either case there is a decrease of vigor on inbreeding and a gain on outcrossing. Wright (1922c) has shown that with the dominance hypothesis the decline in vigor is proportional to the decrease in heterozygosis, regardless of the relative number of dominant and recessive genes and of the degree of dominance. The same decline in vigor with decreasing heterozygosity is true with overdominance.

It is usually impossible in a breeding experiment to differentiate between true overdominance in a pair of alleles, and pseudo-overdominance due to the effects of two pairs of alleles closely linked in the repulsion phase. Only in special circumstances, such as when a mutation has recently occurred in an isogenic stock, can the experimenter be reasonably certain that the effect is due to a single allelic difference. Furthermore, there is the possibility of heterosis due to borderline situations, such as might arise in pseudoalleles with a position effect, which could not even theoretically be classified as due to dominance of linked genes or overdominance. Finally, it should be noted that the various hypotheses may not be equally important in all situations. For example, it is reasonable to expect that overdominance would be more important in determining differences between inbred lines of corn previously selected for general combining ability than in lines not so selected.

If the two hypotheses are not mutually exclusive, neither are they collectively exhaustive. There is no reason to think that multiple factors are any less complex in their interactions than factors concerned with qualitative differences. With the number of genes involved in heterosis, and with the complexity of interactions known to exist in cases where individual gene effects have been isolated and studied, there must surely be all sorts of complex interactions in heterosis. Therefore no single theory can be expected to account for the entire effects of heterosis. Although it is difficult to separate by statistical methods the effects of dominance and epistasis, it may be possible to construct simple models which are of some utility.

DOMINANCE

Davenport (1908) was the first to point out the now well-recognized fact that in most cases the dominant character is beneficial to the organism possessing it, while the recessive has a weakening effect. He noted that this could help explain the degeneration that usually follows inbreeding. Davenport was thinking of relatively few factors with individually large effects, whereas at present, more emphasis is given to multiple factors. But he was close to the ideas now held.

Keeble and Pellew (1910) found that hybrids between two pure varieties of peas were taller than either parent. In this case, two different dominant factors were involved—one resulting in longer internodes and the other in-

creasing their number. Here only two gene pairs were involved, but it was mentioned that similar systems might hold for more complex cases.

A more general development of the dominance hypothesis was given during the same year by Bruce (1910). He designated the frequencies of dominant and recessive alleles as p and q in one breed and P and Q in the other. The array of individuals in the two groups will then be $(p^2DD + 2pqDR + q^2RR)^n$ and $(P^2DD + 2PQDR + Q^2RR)^n$, where D and R are the dominant and recessive alleles and n is the number of factor pairs involved.[1] If these two populations are crossed, the mean number of homozygous recessive loci is nqQ, whereas the average number for the two parent populations is $n(q^2 + Q^2)/2$. The former is the geometric mean of the two parental recessive genotype frequencies while the latter is the arithmetic mean. Since the geometric mean is always less than the arithmetic, the number of homozygous recessive loci will always be less in the hybrid population than the mean number in the two parent populations. If either or both the parent populations are inbred the decrease will be greater.

Bruce then said:

> If, now, it be assumed that dominance is positively correlated with vigor, we have the final result that the crossing of two pure breeds produces a mean vigor greater than the collective mean vigor of the parent breeds. . . . I am aware that there is no experimental evidence to justify the assumption that dominance is correlated with a "blending" character like vigor; but the hypothesis is not an extravagant one, and may pass until a better takes the field.

The average proportion of recessive homozygotes in the parents, which is $(q^2 + Q^2)/2$, may be rewritten as $qQ + (q - Q)^2/2$. This is always larger than qQ, the proportion in the hybrid, unless q and Q are equal. Although Bruce didn't mention this, after one generation of random mating the proportion of recessives in the hybrid population becomes $(q + Q)^2/4 = qQ + (q - Q)^2/4$, which shows that half the gain in vigor is lost as soon as random mating begins.

Bruce concentrated his attention on the decrease of homozygous recessive loci in the hybrid, and postulated a correlation between recessiveness and deleterious effect. He could have used the same algebraic procedures to show that crossing produces an increase in heterozygous loci, and thus based a theory of hybrid vigor on overdominance. He showed remarkable foresight in choosing the former, at a time when he had no evidence of a correlation between dominance and beneficial effect.

1. The notation used by Bruce implies equal frequency of dominant and recessive alleles at all loci. This assumption is not at all necessary for the argument, and I think that what Bruce really meant was

$$\prod_{i=1}^{n} (p_i^2 DD + 2p_i q_i DR + q_i^2 RR) .$$

Objections to the dominance hypothesis were made largely on two grounds. First, if vigor is not a product of heterozygosity as such, it should be possible by selection to obtain individuals which are homozygous for all the beneficial dominant factors, and hence have the same vigor as hybrids. Secondly, in the F_2 of a cross between two inbred strains there should be a skew distribution of the trait being measured—since the dominant and recessive loci would be distributed according to the expansion of $(3/4 + 1/4)^n$, where n is the number of factors.

These objections were largely removed when Jones (1917) pointed out that, with linkage, the consequences of the dominance hypothesis were much closer to those postulating superior heterozygotes. If a detrimental recessive were linked with a favorable dominant, the heterozygous chromosome would be superior to both homozygotes, and the linked combination might not break up readily. Later, Collins (1921) showed that with a large number of factors, regardless of linkage, the skew distribution disappears. The probability of getting all the beneficial dominants into one homozygous strain becomes vanishingly small, so the objections hold only if a small number of factors is assumed.

Most of the mutations known in Drosophila and elsewhere are recessive, and practically all are in some way deleterious. Even if dominant and recessive mutations were occurring with equal frequency, the deleterious mutations in a population at any time would be mostly recessive, since the dominants would be rapidly eliminated. It is to be expected—and it has been often observed—that at most unfixed loci the recessive is deleterious in comparison with its dominant allele.[2]

Almost thirty years ago Sewall Wright (1922c) wrote:

> Given the Mendelian mechanism of heredity, and this more or less perfect correlation between recessiveness and detrimental effect, and all the long-known effects of inbreeding—the frequent appearance of abnormalities, the usual deterioration in size, fertility, and constitutional vigor in the early generations, the absence of such decline in any one or all of these respects in particular cases, and the fixation of type and prepotency attained in later generations—are the consequences to be expected.

It has been shown many times that populations actually contain a large number of detrimental recessives—sufficient to account for a large decline in vigor on inbreeding. In *Drosophila pseudoobscura*, Dobzhansky *et al.* (1942) found that almost every fly examined had at least one concealed lethal. Further evidence that at least some heterosis is due to dominant favorable genes is provided by the experiments of Richey and Sprague (1931) on convergent improvement in corn.

2. I consider the statement that a dominant is beneficial and the statement that a recessive is deleterious as meaning the same thing. Since a geneticist ordinarily can study gene effects only by substituting one allele for the other, he cannot distinguish what each factor is doing individually or whether it is harmful or beneficial except relative to its allele. That is, he can only tell what the effect of the substitution is.

OVERDOMINANCE

The concept of a stimulating effect of hybridization began independently with Shull (1908, 1911b) and East (1908). It was assumed that there was a physiological stimulus to development which increased with the diversity of the uniting gametes—with increasing heterozygosis. East (1936) elaborated the idea further by postulating a series of alleles each having positive action functions, and with these functions to some extent cumulative. As the alleles became more and more divergent in function, the action was postulated to become more nearly additive in the heterozygote.

At the time when East and Shull first formulated the hypothesis, there was no direct evidence of any locus at which the heterozygote exceeded either homozygote. For a number of years, overdominance as an explanation of heterosis largely was given up because of the failure to find such loci.

Stadler (1939) pointed out that in certain of the *R* alleles in corn a situation obtains in which certain heterozygotes have more areas pigmented than either homozygote. He suggested that genes acting in this manner could result in overdominance for such characters as size and yield. Other such loci are known in corn.

There are now several cases in the literature of single genes with heterotic effects. In most of these it is not possible to rule out the possibility of close linkages giving pseudo-overdominant effects. In particular, many cases may turn out to be pseudoallelism, but the consequences for the animal or plant breeder would not be changed.

Several workers (Teissier, 1942a; Robertson, unpublished) have found persistent lethals in Drosophila population cage experiments. If these are not due to individually heterotic loci, extremely close linkage must be postulated. Also certain recessive genes, such as *ebony*, come to an equilibrium with their normal alleles in population cages. One of the most convincing cases is that of the eye color mutant described by Buzzati-Traverso in this volume. This mutant persists in the population, and was found in three independent stocks. It is quite improbable that in each of these cases the gene happened to be linked in the repulsion phase with another harmful recessive.

The idea of superior heterozygotes has been upheld by Hull (1945) who suggested the word overdominance. Hull's original argument for overdominance is a simple one. He noted that in most cases the hybrid between two inbred maize lines has a greater yield than the sum of the two inbreds. This would not be possible with dominant genes acting in a completely additive manner—unless it were assumed that a plant with no favorable dominants had a negative yield.

The validity of this argument depends on the unimportance of epistasis in corn yields. Evidence on this point is very incomplete and somewhat contradictory. Neal (1935) reported that the F_2 yields were almost exactly intermediate between the F_1 and the average of the parents. This would suggest that

epistatis is not important or else that there is some sort of cancelling out of various effects. On the other hand, Stringfield (1950) found that in many cases backcrosses showed consistently higher yields than the F_2. This suggests some sort of interaction, as if some of the gene combinations selected for during the inbreeding process were active in the backcross, but were broken in the F_2. None of these data give any evidence as to the importance of epistasis in determining the difference between an inbred line and a hypothetical line with none of the favorable dominants, since the data do not extend into this range. It is in this range where non-additivity might be expected to be most pronounced.

Hull's second argument is based on results obtained by the technique of constant parent regression. The regression of F_1 on one parent, with the other parent held constant, has different expectations when there is overdominance than when there is dominance. With overdominance the regression may be negative when the constant parent is high-yielding, so the regression surface is different from that expected with dominance. In this volume Hull gives data which conform with this expectation.

Overdominance is not the only possible explanation of such results, as Hull has pointed out. In addition, the constant parent regression technique, or any technique making use of yield data on inbred lines, is complicated by the difficulty of obtaining consistent results with inbreds. Another possibility is that the factors responsible for yield in inbreds are largely different genes from those determining the yield in the hybrids. This possibility will be considered later.

For these reasons it is still not possible to be sure of the importance of overdominance from Hull's methods. They are at least strongly suggestive, and recent data from Robinson *et al.* (1949), obtained by an entirely different procedure, also gave evidence of overdominance.

MAXIMUM HETEROSIS WITH THE DOMINANCE HYPOTHESIS

In this discussion several assumptions are made. Most of these have been implicit in most discussions of heterosis, but it is best that they be clearly set forth at the outset. The assumptions are:

1. Genes concerned with vigor are dominant, and in each case the dominant allele is beneficial and the recessive deleterious. This is an assumption of convenience which does not alter the essential nature of the hypothesis. The conclusions still hold if dominance is not complete. Also there are loci in which the recessive is advantageous or in which the heterozygote is intermediate; but these are of no consequence for heterosis and therefore can be omitted from the discussion.

2. There is complete additivity of effects between loci—no epistasis.

3. There are no barriers to recombination that prevent each gene from reaching its own equilibrium frequency independently of other loci.

4. The gene and phenotype frequencies of the parent population are at their equilibrium values.

5. Increased vigor results in, and can be measured in terms of, increased selective advantage, though the selection may be natural or artificial. This assumption restricts the discussion to those cases in which heterosis results in changes in the same direction as selection had previously been acting. Such an assumption appears to be valid for yield characters in field crops, and for viability and fertility as is measured in Drosophila population studies. It is highly questionable for such things as increase in size of hybrids between wild varieties or species, where natural selection pressure may well have been toward an intermediate size.

Under this assumption the increase of vigor on hybridization depends directly on the number of loci which are homozygous recessives in the parent, but which become heterozygous in the hybrid. The individual or population of maximum vigor is one in which every allelic pair contains at least one dominant. The actual attainable heterosis would be less than this in any particular case.

Consider the case of complete dominance. The recessive phenotype is assumed to have a selective disadvantage of s. That is, the dominant and recessive phenotypes are surviving and reproducing in the ratio of 1 to $1 - s$. The rate of mutation from A to a is u per gene per generation. Reverse mutation will be ignored as it can be shown to have a negligible effect on the equilibrium gene frequency attained.

Genotype	AA	Aa	aa
Frequency	P	$2Q$	R
Selective value	1	1	$1 - s$

$$P + 2Q + R = 1$$

Under these assumptions, the frequency of gene A will be $P + Q$, while the frequency of a will be $Q + R$. With mutation from A to a at rate u, the frequency of A will be reduced in one generation by $u(P + Q)$ and the frequency of a increased by the same amount. Likewise, due to the effect of selection, the frequency of a will be decreased by sR. Therefore the gene ratio, $(P + Q)/(Q + R)$, will change in one generation due to the effects of mutation and selection to

$$\frac{(P+Q)(1-u)}{(P+Q)u+Q+R-sR}.$$

When equilibrium is reached the gene frequency will no longer change from generation to generation which, stated algebraically, is

$$\frac{P+Q}{Q+R} = \frac{(P+Q)(1-u)}{(P+Q)u+Q+R-sR}.$$

This has the solution, $R = u/s$. (For a more pedantic demonstration of this, see Crow, 1948.)

The average reduction in selective value of the population due to a detrimental factor will be the product of the selective disadvantage of the factor and the proportion of individuals possessing the factor. This amounts to (s) (u/s), or, simply, u, the mutation rate. Hence, the effect of a detrimental gene on the selective value of the population is equal to the mutation rate to that gene, and is independent of the selective disadvantage which that factor causes, as was first pointed out by Haldane (1937). This fact, which at first appears paradoxical, is readily understandable when one notes that a mildly deleterious mutant persists much longer in the population, and hence affects many more individuals than one which has a greater harmful effect.

The total effect on the population of all the loci capable of mutating to deleterious recessives is simply the sum of the individual mutation rates as long as the gene effects are additive. If there are n such loci with an average mutation rate of $\bar{u}$, the net reduction in selective value due to all homozygous detrimental recessives at all loci in which they occur is $n\bar{u}$. This is also approximately correct if the factors are multiplicative, provided the individual effects are small.

The product $n\bar{u}$ is probably in the vicinity of .05 (Crow, 1948). This means that if all the deleterious recessives were replaced by their dominant alleles, the selective advantage of an equilibrium population would be increased by about this amount. This could be considered as the maximum average improvement in vigor, as measured in terms of selective advantage, that could occur due to hybridization. This means that the dominance hypothesis cannot, under the conditions postulated, account for average increases of more than a few per cent in vigor.

There are several reasons why the 5 per cent figure given above may be too large. One is that many deleterious factors considered to be recessive may not be completely recessive. Stern and Novitski (1948) and Muller (1950) have shown that the majority of lethals and detrimentals that occur in laboratory cultures of Drosophila are not completely recessive. Even if the detrimental effect of the heterozygote is much less than that of the homozygote, the greatest selection effect will still be on heterozygotes because of their much greater frequency in the population. Thus, from the population standpoint, these factors would be acting more like dominants than recessives. This means that each locus would have a detrimental effect of $2u$ instead of u (since a dominant gene would be responsible for twice as many "genetic deaths" as a recessive), but the locus would be unimportant for heterosis. Since the n in the formula refers only to the number of loci which are capable of mutating to a completely recessive allele, its value may be smaller than previously assumed and the product $n\bar{u}$ proportionately less.

It has been assumed that the parent populations are at equilibrium be-

tween selection and mutation pressures. This assumption probably is not strictly correct for any population. Any equilibrium involving occurrences as rare as mutations must be slow of attainment. Hence many if not most populations must not be at equilibrium. Probably the most common way in which a population gets out of equilibrium is by an alteration of the breeding structure or population number so that the effective amount of inbreeding is changed. If the change in population structure is such as to increase the amount of homozygosity, a new equilibrium is reached comparatively rapidly through the elimination by selection of the recessives which have been made homozygous. On the other hand, if the change in population is such as to decrease the amount of homozygosity a new equilibrium is attained only through the accumulation of new mutations. This is an extremely slow process.

Since the return to equilibrium is much slower when the population changes in the direction of less inbreeding, it follows that most populations which are out of equilibrium will be out in the direction of having too few detrimental recessives. Therefore the effect of fluctuations in population size and breeding structure will be on the average such as to increase the fitness of the population. For this reason, the average loss of fitness per locus is probably less than the mutation rate. Fisher (1949) has pointed out that if the yield of a crop is near a "ceiling," the relative effect of each factor conditioning yield becomes less. There will be a similar tendency for the population to be out of equilibrium because of the slowness of occurrence of the mutations required to bring the population to the new equilibrium level.

Another factor also pointed out by Fisher is that complete lethals and highly deleterious factors contribute to the mutation rate but, at least in grain crops, have no appreciable effect on yield since they are crowded out by other plants.

All of these factors make the 5 per cent figure an overestimate, so it should be regarded as a maximum. The true value may be much less. In this connection Fisher (1949) said:

> . . . it would appear that the total elimination of deleterious recessives would make less difference to the yield of cross-bred commercial crops than the total mutation rate would suggest. Perhaps no more than a 1 per cent improvement could be looked for from this cause. Differences of the order of 20 per cent remain to be explained.

These considerations make it difficult to explain, in terms of the dominance hypothesis, cases in which two equilibrium populations produce hybrids with considerable heterosis, or in which crosses between inbred lines average appreciably more than the randomly mating populations from which they were derived.

This discussion is relevant only when the character is measurable in terms of selective value. For yield characters subject to any high degree of artificial selection an increase in yield is probably accompanied by a greater propor-

tional increase in selective value. Thus any conclusions about maximum proportional increase in selective value would hold *a fortiori* for yield. Fisher (1949) reaches a similar conclusion when he says: "If the chance of survival is equated to the yield, as is reasonable with grain crops."

Another assumption is that the hybrids are compared with equilibrium populations. There is room for question, particularly with domestic plants and animals, as to whether selection has been occurring long enough and its direction has been consistent enough for a gene frequency equilibrium to have been attained. Another point that must be remembered in discussions of maize is that commercial hybrids are not random combinations of inbred lines, but highly select combinations. An *average* hybrid may have a yield very close to that of a randomly mating population. Thus the argument of this section may not be relevant for corn. But it can hardly be true that the high yield of certain corn hybrids is due to the elimination of deleterious recessives during inbreeding.

The quantitative limit placed on average improvement on hybridization with the dominance hypothesis does not hold for overdominant loci. A locus at which the homozygote AA has a selective disadvantage of s with respect to the heterozygote, and the homozygote $A'A'$ has a disadvantage of t, will come to equilibrium with gene frequency of A equal to $t/(s + t)$, and the frequency of A' equal to $s/(s + t)$ (Wright, 1931b; Crow, 1948). The average reduction in selective advantage of the population due to the two homozygous genotypes comes out to be $st/(s + t)$. The loss in fitness of the population is of the order of magnitude of the selection coefficients, as Haldane (1937) has first shown, whereas with a detrimental recessive, the loss is of the order of the mutation rate. Hence a single overdominant locus has a tremendously greater effect on the population fitness than a single locus with dominance or intermediate heterozygote. If such loci are at all frequent they must be important. The question is: how frequent are they?

Even with overdominance it is difficult to understand large average increases in selective advantage of hybrids between equilibrium populations. Such populations should be somewhere near their optimum gene frequencies, which means that the hybrids would be about the same as the parents. It may be that, on the average, hybrids do not greatly exceed their parents in selective advantage, and that the cases of increased size observed in variety crosses and occasionally in species crosses are nothing but *luxuriance*. If so, they are much less difficult to explain.

As Bruce showed in 1910, if the parents differ at all in gene frequencies, the hybrids will be more heterozygous. If both parents are at equilibrium they should have, for additive genes, approximately the same frequencies. But what differences there are—due to chance, for example—will amount to much more in an overdominant than in a dominant locus because the former has a gene frequency much nearer .5.

POPULATION VARIANCE

The same considerations which show that an overdominant locus has a much greater effect on average population fitness than a dominant locus also show that an overdominant locus has a much greater effect on the population variance. If the selective values of the three genotypes, AA, Aa, and aa are 1, 1, and $1 - s$ respectively, the frequency of aa genotypes is u/s and the average selective value $1 - u$. The variance in fitness will be su. On the other hand, with an overdominant locus where the fitnesses of the three genotypes are $1 - s$, 1, and $1 - s$, the mean fitness is $1 - s/2$. The variance in fitness is $s^2/4$.

The ratio of these variances is $s/4u$, which means that an overdominant locus causes a population variance $s/4u$ times as great as that resulting from a recessive locus of the same selective disadvantage. If $4u$ is 10^{-5}, this amounts to 100 for $s = .001$, or is 1000 for $s = .01$. This makes an overdominant locus with these selective values equivalent to 100 or 1000 ordinary loci in its effect on the population variance. Haldane (1950) has emphasized the importance of loci with adaptively superior heterozygotes in increasing the variance of natural populations.

From this we must conclude that there doesn't have to be a very high proportion of overdominant loci for overdominance to be the most important factor in the genetic variance of the population. If much of the genetic variance of a population is due to overdominance, this would explain the great slowness of selection. Characters with high genetic determination but low parent-offspring correlation might be due to this cause.

The facts of hybrid corn also are consistent with this. Ordinary selection has not been effective. Yet there is a great deal of variation in an open-pollinated variety. It has been relatively easy to find combinations of inbred lines that have yields well above the open-pollinated averages. There appears to be a relatively high degree of genetic determination of yield, but relatively low heritability. These results are not impossible with dominant genes, especially with epistasis, but are precisely what would be expected if some of the variance were due to overdominant loci.

A population with many overdominant loci is always well below its maximum possible fitness. It is expected that such factors could eventually be replaced in long evolutionary periods. This might occur by an appropriate mutation, by duplication, or by modifiers. Or a population with too many overdominant loci might disappear due to inter-population competition. But at any particular time, a population may have a small proportion of such loci, and it does not require many for these to be the major source of variation.

DO THE SAME GENES DETERMINE VARIATION IN INBREDS AND HYBRIDS?

The rarer a recessive phenotype is in a population, the greater will be its relative increase in frequency on inbreeding. If the frequency of the recessive

gene is q, the frequency of recessive homozygotes in a randomly mating population is q^2. With increasing amounts of inbreeding, the frequency changes from q^2 to q. The smaller the value of q, the greater is the ratio of q to q^2. If a gene is highly deleterious it will be very rare in the population. Hence the genotypes which are most deleterious are those which have the greatest relative increase in frequency on inbreeding.

These relationships are brought out in the following figures, based on a mutation rate of 10^{-6}. The ratio given is the ratio of homozygous recessives in a homozygous population as compared with one which is mating at random.

Selective disadvantage (s)	.0001	.001	.01	.1	lethal
Gene frequency (q)	.1	.032	.01	.003	.001
Ratio (q/q^2)	10	32	100	316	1000

This means that highly deleterious recessives, which ordinarily have an effect on the population only of the order of the mutation rate, become much more important with inbreeding and may become the major factors in determining the fitness of an inbred population. This might to some extent be offset by selection during the inbreeding process, but such selection would be directed against factors which are of no consequence in a more heterozygous population.

The detrimental recessive factors referred to here include the lethals and semilethals (such as chlorophyll deficiencies) that show up during inbreeding. But more important are the larger number of factors, not individually detectable, which collectively result in the loss of vigor with inbreeding despite rigorous selection.

On the other hand, the major part of the variance of a non-inbred population may well be determined by genes of intermediate frequencies, from .1 to .9. The effect of such factors in determining the population variance in fitness would change only slightly with inbreeding.

As an example, consider a hypothetical population mating at random whose variance is made up of two components. Ninety per cent of the variance is due to relatively common loci with gene frequencies of the order of .5. The other 10 per cent is due to loci with recessive gene frequencies of the order of .01 or less. Now when this population is inbred without selection, the variance due to the common genes will not change greatly but the variance due to the recessive loci will increase by a hundred fold or more. Thus the factors which originally contributed only 10 per cent to the variance may now contribute over 90 per cent of the variance between the various inbred lines derived from the population.

Gene frequencies of the order of .5 might result from several causes. They might be genes which are advantageous in one geographical location and disadvantageous in another so as to form a cline. Or there might be seasonal differences in selective value. They may be due to complex interactions with other loci or be of extremely small selective advantage or disadvantage. But

another explanation is selective superiority of heterozygotes (Haldane, 1950), at least for those factors of importance in heterosis.

If yield is determined entirely by dominant factors, the correlation between inbreds and their hybrids should be positive. If it is due to overdominant loci, the correlation should be generally positive, though there would be negative correlations between yield of hybrids and inbreds when the other inbred is constant and high yielding. If both factors are involved and overdominant loci are relatively important in hybrids while dominants are important in inbreds, the correlation would approach zero. The experience of corn breeders has been that selection for yield during inbreeding is relatively ineffective, and that the correlation of hybrid with inbred yield, though positive, is small.

With overdominant loci the effect of a certain percentage increase in heterozygosity is to cause the vigor to increase by a certain amount. Decreasing the heterozygosity by the same percentage would cause a decrease of approximately the same amount. On the other hand, with dominant loci, making the original equilibrium population more heterozygous would cause a very slight increase, whereas making the population more homozygous would have a decreasing effect of a much greater amount. Therefore it is easier to account for inbreeding depression by dominant loci than to account for increase in vigor on hybridization above the level of a random mating population.

I should like to suggest the following interpretation of the effects of inbreeding and hybridization: The deleterious effects of inbreeding and the recovery on hybridization are mainly due to loci where the dominant is favorable and the recessive allele so rare as to be of negligible importance in a non-inbred population. Variance of a non-inbred population, and hybrid vigor when measured as an increase over an equilibrium population, are determined largely by genes of intermediate frequency, probably mostly overdominants.

OVERDOMINANCE AND GENE ACTION

In order to have overdominance it is not necessary that the immediate gene products of the heterozygote exceed in quantity or variety those of either homozygote. At the level of the immediate gene product, or any intermediate state, the effect of the heterozygote may be intermediate between the two homozygotes and still result in a greater final result. Any kind of situation in which something is produced for which an intermediate amount is optimum could be such that the heterozygote is nearer this optimum than either homozygote.

A model for such cases is found in the sulfanilamide-requiring strain of Neurospora reported by Emerson (1948). When this mutant is present the heterokaryotic state of the suppressor gene results in more nearly the opti-

mum amount of para-amino benzoic acid than either homokaryon. Other cases, less known biochemically, may be similar.

I think that it is doubtful whether such a system would persist for long evolutionary periods. Alleles of intermediate productivity could arise and replace the originals. Also modifiers altering the expression of the homozygotes would have considerable selection pressure. Or if the alleles were antimorphic, the situation might be resolved by duplication, as Haldane (1937) has suggested. It is significant that the system reported by Emerson is not one which is ordinarily of importance, but acts only in the presence of the sulfanilamide-requiring mutant.

A form of gene action that appears more likely to account for instances of overdominance is one in which the two alleles differ qualitatively or each does something that the other fails to do. Instances of mosaic dominance provide excellent examples. This has been demonstrated for the *scute* series of bristle characters in Drosophila and for color pattern in beetles (Tan, 1946). Other examples are provided by the *A* and *R* loci in maize.

Similar examples of *physiological mosaic dominance* are found where the heterozygote apparently produces something approximating—at least qualitatively—the total effect of the two homozygotes. An example is rust resistance in flax, where each strain is resistant to a certain rust but the hybrid is resistant to both (Flor, 1947). By the usual tests for allelism, the two resistance factors are alleles. Another series of examples is found in the blood group antigens in man, cattle, and elsewhere. In almost every instance the heterozygote has all the antigenic properties of both homozygotes (Irwin, 1947). The presence of both the normal and abnormal types of hemoglobin in humans heterozygous for the gene for sicklemia provides another example (Pauling, 1950).

Many instances of overdominance may have a similar explanation. This is the kind of action that East (1936) postulated in his discussion of heterosis due to cumulative action of divergent alleles. It is not necessary that the effects be completely cumulative; only that the net effect on the phenotype be greater in the heterozygote than in the homozygote. Any system in which the alleles act on different substrates to produce the same or different products, or convert the same substrate into different products—neomorphs, in Muller's terminology—could result in overdominance.

Any of the examples listed above may turn out to be closely linked genes (pseudoalleles) rather than alleles. In most cases it is impossible to distinguish between these alternatives. If the overdominance effect is due to linked genes, eventually a crossover should result in a situation where the desirable effects could be obtained in a homozygous individual. If there are position effects, it may be that no homozygous arrangement is as advantageous as one which is heterozygous. Unless there are position effects, it does not seem likely that heterosis due to pseudoallelism would persist for any

great length of time, but in any particular population such factors might be important.

IS INCREASED SIZE ADAPTIVE?

The foregoing arguments are based on the assumption that heterosis is measurable in terms of increased selective advantage. The selection may be natural or man-imposed. This assumption would appear to be reasonable for such factors as fertility and resistance to disease. It also would apply to increase in size or yield, if the direction of selection in the past were in this direction, as in corn. However, it is questionable whether the increase in size that is sometimes observed in variety hybrids is really adaptive.

Mather (1943) and especially Dobzhansky (1950) have emphasized that increased size does not necessarily result in increased fitness in natural populations. Dobzhansky proposed the words *euheterosis* and *luxuriance*, respectively, for increased selective advantage and for mere non-adaptive increase in size. In these terms this discussion has dealt entirely with euheterosis.

If euheterosis occurs in species or variety crosses, it is very difficult to explain. It raises the troublesome question: How can the hybrid between two well adapted strains be better adapted than its parents when there has been no selection in the past for its adaptation? It may be that euheterosis is developed only under some form of selection, as in the inversion heterozygotes studied by Dobzhansky, or in the series of hybrids between inbred lines of corn selected for combining ability.

If large size is not advantageous, luxuriance may be due to the covering of recessive factors which were acting as size bottlenecks and had been selected into the population because of this. Each of the parents might have its growth limited by or held in check by a series of factors, and if some of these were recessive, increased size would be found in the hybrids.

SUMMARY

Since the earliest attempts to explain hybrid vigor in Mendelian terms there have been two principal hypotheses. The first of these is the dominance hypothesis. This notes the observed correlation between recessiveness and detrimental effect and attributes the increased vigor of heterozygosity to the covering of deleterious recessive factors by their dominant alleles. The alternative hypothesis, the overdominance hypothesis, assumes that heterozygosity *per se* is important—that there exist loci at which the heterozygote is superior to either homozygote.

It is clear that the dominance hypothesis is adequate to explain the deterioration that results from inbreeding and the recovery of vigor on outcrossing, but it is difficult to explain how the hybrids could greatly exceed in fitness the equilibrium populations from which their parents were derived. The overdominance hypothesis demands the assumption of a kind of gene

action known to be rare, but it is pointed out that if only a small proportion of the loci are of this type, these may nevertheless be the major factor in the population variance.

The following interpretation is suggested: Inbreeding depression and recovery on crossing are mainly the result of loci at which the favorable allele is dominant and the recessives are at low frequency. On the other hand the variance of heterozygous populations and the differences between different hybrids are due mainly to loci with intermediate gene frequencies. It appears likely that such loci are due to selectively superior heterozygotes, but there are several other possibilities.