

NOTES AND COMMENTS

HETEROZYGOTE SUPERIORITY, SELECTION INTENSITY AND PLATEAURING

J. H. VAN DER VEEN

*Department of Genetics, Agricultural University, Wageningen,
The Netherlands*

Received 7.vii.60

When selecting for a metric character in a population where the best genotype conceivable is a homozygote, the theoretical limit to selective advance will generally not be reached, owing to chance fixation of undesirable alleles and to genetic homeostasis (inertia). The latter implies that other characters such as viability and fertility are affected via linkage or pleiotropy, or that the release of variability from residual blocks of closely linked plus and minus genes is very slow. Operationally these chromosome segments behave as single units in transmission, and if the plus genes are preponderantly dominant, then the heterozygote (*e.g.* inversion heterozygote) is heterotic in the direction of selection and acts towards conservation of genetic variability. Both this kind of overdominance and some types of non-allelic interaction have been invoked in explaining heterosis. With these points in mind, the type of population to be considered here is one in which certain heterozygotes are preferred to any homozygote arising from the gene pool in other than negligible quantities.

For the sake of simplicity it will be assumed that (1) environmental variation is absent, (2) after mass selection random mating is practised, (3) random drift can be neglected, and (4) no selection forces are acting but the one intentionally applied to the metric character, which implies that suspension of selection will not result in change of gene frequencies.

With mass selection in the population defined above, the limit to selective advance is at gene frequencies which maximize the population mean. However, whether this maximum is actually attained depends—given the same assumptions as above—on selection intensity, *i.e.* on the fraction of the population permitted to reproduce itself. In fact, the plateau ultimately reached is a direct function of selection intensity, and may diverge considerably from the maximum attainable. It seems that little or no attention has been paid to this type of artificially induced plateauring. The following paragraphs present a detailed analysis of the phenomenon for one allelic pair.

Let p be the relative frequency of the plus allele A , and let $p+q = 1$. Then on random mating the relative frequencies of AA , Aa , and aa , are p^2 , $2pq$ and q^2 , respectively. The genotypic values can be denoted by $m+d$, $m+h$ and $m-d$, respectively. In our model $h > d$ since $Aa > AA > aa$. The population mean is $m + (2p-1)d + 2p(1-p)h$, which for $h > d$ has its maximum at $p_m = (d+h)/2h$. Any change of p towards p_m brings about an increase of the mean. In passing it may be noted that the additive variance, which is the part of the genotypic variance explained by regression on gene dosage, is $2p(1-p)[d - (2p-1)h]^2$, which becomes zero at p_m .

At the same time heritability in the narrow sense becomes zero, since its numerator is the additive variance.

In breeding programmes and selection experiments, selection intensity (v) is often more or less constant over successive cycles. If $v > 2pq$, the preferred Aa individuals are supplemented by AA , and if $v > p^2 + 2pq$, even by aa individuals. Any change in p brings about a change in the relative frequencies of the genotypes. It is therefore futile to describe response to artificial selection by constant parameters for relative selective advantage.

Now it can be shown (Appendix I) that mass selection followed by random mating among the selected individuals leads to a stable equilibrium at $p = v$ for $v \geq \frac{1}{2}$, or at $p = \frac{1}{2}$ for $v \leq \frac{1}{2}$. Therefore, if v is larger or smaller than p_m , the ultimate plateau will fall short of the maximum population mean. For example, $p_m = 0.875$ for $h/d = \frac{2}{3}$, and the maximum will be approached closer by relaxed selection ($v = 0.75$) than by intense selection ($v = 0.25$). At first sight it may seem even more paradoxical that the disadvantage of intense selection is weakened by environmental variation, which causes a discrepancy between the v best genotypes and the v best phenotypes. Assuming, as before, absence of environmental variation, it will be found (Appendix II) that for $p < p_m$, maximum progress per cycle is achieved with $v = p^2 + 2pq$. This amounts to complete selection against aa .

It would be interesting to explore other systems of selection. When choosing for further breeding the progeny of parents individually selected after random mating, the same equilibrium conditions as above will be obtained (proof not given here). In judging the breeding value of a parent on the basis of its offspring from random mating, it can easily be verified that in the case of $h > d$ the conclusion will be that AA is better than Aa , and Aa is better than aa when $p < p_m$, and the reverse is the case when $p > p_m$. This will lead to certain oscillation patterns of p , again dependent on v .

Even when more elegant algebraic methods are found, this type of problem requires a good deal of inductive reasoning. In models with more loci an empirical approach by means of computers may become necessary. Considering two unlinked loci, the relation $AaBb$ is better than $AABB$, and $AABB$ is better than "other genotypes", requires non-allelic interaction and cannot be mimicked by overdominance. Here an equilibrium exists at the diheterozygote frequency $2v(1-2v)$; see Appendix III. It is tentatively suggested that starting from any distribution of genotypes the following stable equilibria will be reached. All genotypes $AABB$ for $v \geq \frac{1}{2}$; all selected parents $AaBb$ for $v \leq \frac{1}{4}$; and $p_a = p_b = 2v$ for $\frac{1}{2} \geq v \geq \frac{1}{4}$.

In conclusion it may be argued that especially in populations at higher breeding levels one should be prepared to meet the above-mentioned type of induced plateauing, since certain loci and gene blocks connected with heterozygote superiority resist fixation and thus represent an increasing part of the total genotypic variation.

Appendix I

In selecting a fraction v of the best genotypes from a one locus population after random mating, three relations can occur.

Relation 1. $v \leq 2p_0q_0$. This implies at least $v \leq \frac{1}{2}$, and leads to $p_1 = \frac{1}{2}$ and $2p_1q_1 = \frac{1}{2}$. Consequently $p_2 = \frac{1}{2}$, $p_3 = \frac{1}{2}$, etc.

Relation 2. $p^2_0 + 2p_0q_0 \geq v > 2p_0q_0$. Here $q_1 = p_0q_0/v$, and $p_1 = (v - p_0q_0)/v$. From $v > 2p_0q_0$ follows $q_1 = p_0q_0/v < \frac{1}{2}$, and consequently $p_1 > \frac{1}{2}$. Therefore, in the following, one needs to consider only the case $p > \frac{1}{2}$. From $p_1 - p_0 = (v - p_0)q_0/v$ it is seen that $v > p_0$ implies $p_1 > p_0$, and that $v < p_0$ implies $p_1 < p_0$. Furthermore, $p_1 > v > p_0$ has been excluded, since after substituting $v = p_0 + \epsilon$ in $(v - p_0q_0)/v > v$, one obtains $1 - 2p_0 > \epsilon$, which implies $p_0 < \frac{1}{2}$. Finally, $p_0 > v > p_1$ is impossible, as substitution of $p_0 = v + \epsilon$ in $(v - p_0q_0)/v < v$ leads to $1 - 2v > \epsilon$. This means $v < \frac{1}{2}$, and $p_1 < \frac{1}{2}$, which has been excluded. One can now distinguish three cases.

- (a) $v > p_1 > p_0 > \frac{1}{2}$ for $v > \frac{1}{2}$ and $v > p_0$,
- (b) $p_0 > p_1 > v \geq \frac{1}{2}$ for $v \geq \frac{1}{2}$ and $v < p_0$,
- (c) $p_0 > p_1 > \frac{1}{2} \geq v$ for $v \leq \frac{1}{2}$ and $v < p_0$.

The direct conclusion is that in the cases (a) and (b) p moves monotonically in the direction of v , leading to a stable equilibrium at $p = v$. In case (c) p moves monotonically towards $\frac{1}{2}$, during which process one obtains $v < 2pq$, where the stable equilibrium of "relation 1" is reached. It is interesting to note that if $v = 1 - p_0 > \frac{1}{2}$, then $p_1 = (1 - p_0 - p_0 + p^2_0)/(1 - p_0) = v$, which is the equilibrium condition.

Relation 3. $v > p^2_0 + 2p_0q_0$. Here $p_1 = (p^2_0 + p_0q_0)/v = p_0/v$. Therefore $p_1 > p_0$, which for $v > \frac{1}{2}$ ultimately leads to "relation 2", and for $v < \frac{1}{2}$ to "relation 2" or "relation 1".

Appendix II

Let p_m be the value of p maximizing the population mean, and let $p < p_m$. Then maximum progress per selection cycle is obtained in "relation 3" ($p_1 = p_0/v$) for the smallest v value, that is for $v = p^2_0 + 2p_0q_0$, and in "relation 2" ($p_1 = 1 - p_0q_0/v$) for the largest value of v , which is also $v = p^2_0 + 2p_0q_0$. Finally, $v < 2p_0q_0$ implies $p_1 = \frac{1}{2}$, so for these values of v maximum progress is not achieved, as $p_1 > \frac{1}{2}$ for $v > 2p_0q_0$ (Appendix I). With v for maximum progress one has $p_1 = 1 - p_0q_0/v = 1 - p_0q_0/(p^2_0 + 2p_0q_0) = 1/(1 + q_0)$.

Appendix III

Let $f_1, f_{11}, f_{111}, \dots$ be the relative frequencies of the heterozygotes $Aa, AaBb, AaBbCc, \dots$. At equilibrium $p_a = v$, so $f_1 = 2p_aq_a = 2v(1 - v)$. If v takes out $AaBb$ and $AABB$ individuals only, it can be similarly calculated for $AaBb > AABB > \dots$ "all other genotypes", that an equilibrium exists at $p_a = p_b = 2v$, and $f_{11} = 2v(1 - 2v)$. Random mating equilibrium between loci has not been assumed. In general, in the case of n unlinked loci, one finds for $AaBb \dots Zz > AABB \dots ZZ > \dots$ "all other genotypes", that $p_a = p_b = \dots = p_z = 2^{n-1}v$, and $f_{11 \dots 1} = 2v(1 - 2^{n-1}v)$.