ABSTRACT - Epistatic effects involving genic combinations of fixed and non fixed genes are shown to contribute to the genotypic mean of any population. These effects define specific additive x additive and additive x dominant epistatic components. As such components are not estimable, their relative importance cannot be assessed. These epistatic effects can cause bias in the estimates of the additive and dominance components to which they are confounded. The magnitude of the bias depends on the relative values of the epistatic effects, comparatively to deviations d and h, type of prevailing epistasis and direction of dominance.

Index terms: biometrical genetics, components of means.

INTRODUCTION

The generation mean analysis is commonly employed in studies of inheritance of quantitative traits. In relation to yield and its components in Indian mustard, Rishpal & Kumar (1993) verified duplicate epistasis and markedly genic effect x environment interaction. Both additive and non additive
genetic components were found to be important for traits expression. Rahman et al. (1994) studied the genetic control of seedling root characteristics in maize. The analyses showed that dominance effects were greater than additive genetic effects, and epistasis was an important component of the genotypic values of individuals. Ramsay et al. (1994) combined generation mean analysis and triple test cross to study the inheritance of yield and its components in swede. In relation to direction of dominance, reciprocal effects and epistasis, relevant differences between the polygenic systems were evidenced. In rice, Mgonja et al. (1994) and Saha Ray et al. (1994) obtained evidence of duplicate epistasis in the polygenic systems responsible for stem and mesocotyl length, respectively. Both additive and non additive genic effects were important.

Das & Griffey (1995) studied the genetic control of wheat adult-plant resistance to powdery mildew. Additive effects were predominant in all six crosses. In relation to five crosses there was evidence of digenic epistasis. Dominance effects were significant in four crosses. Barakat (1996) employed generation mean analysis in combination with an analysis of variation to study the inheritance of wheat characters related to immature embryo regeneration capacity. The author found that epistatic effects were more important than either additive or dominance genic effects to determine the five in vitro traits.

Balatero et al. (1995) did not verify epistatic effects determining the androgenetic response in hexaploid triticale. The results showed that dominance was an important cause of high embryo induction. Absence of epistasis was also verified by Gingera et al. (1995), in a study of inheritance of delayed first pustule appearance to common leaf rust in sweet corn, and Holtom et al. (1995), in relation to maturity and seed characters in sunflower.

As seen, this methodology provides information on the relative importance of average effects of the genes (additive effects), dominance deviations, and effects due to non allelic genic interactions, in determining genotypic values of the individuals and, consequently, mean genotypic values of families and generations. It is interesting to note that in most of the papers there was evidence of epistasis. Thus, in genetic studies assuming additive-dominance model, the estimates of the linear components of means can be biased due to epistatic effects. This has been shown for the case of polygenic system without fixed genes. This work was carried out to assess the consequences of epistasis in polygenic systems with fixed genes and is an extension of part of the paper of Pooni & Treharne (1994), in which the problem was first considered.

COMPONENTS OF MEANS

Take a polygenic system composed of \( k \) genes with independent assortment and two homozygous parents, which are different in relation to \( k_1 \) of the \( k \) genes \( (k_1 \leq k) \). The others \( k - k_1 = k_2 \) genes are fixed in the parents. Regardless of the distribution of the non fixed genes in the parents, the \( F_{\infty} \) population, obtained after an infinitely large number of generations of selfing without selection, mutation, migration and genetic drift, will be made up of \( 2k_1 \) pure...
lines, all with frequency \((1/2)^k\). Thus, allowing for interaction between genic combinations of two loci:

\[
F = \sum_{i=1}^{k_1} m_i + \sum_{i=1}^{k_2} d_i + \sum_{i=1}^{k_3} i_n - \sum_{i=1}^{k_3} i_n + \sum_{i=1}^{k_4} i_n = m
\]

where \(k_1\) and \(k_2\) are the numbers of fixed genes that increase and decrease trait expression, respectively \((k_1 + k_2 = k_2)\).

For all loci the parameter \(m\) is the mean of the genotypic values of the homozygotes; \(d\) is the deviation between the genotypic value of the homozygote of greater expression and \(m\); \(i\) is the epistatic effect due to interaction between two homozygous genic combinations (Mather & Jinks, 1974; Kearsey & Pooni, 1996). Therefore, the mean of the \(F_0\) generation depends on the genotypic means of the loci with non fixed genes, the genotypic values of the loci with fixed genes, and the epistatic effects between fixed genic combinations. Obviously, there is maximum flexibility for selection of superior genotypes in segregant generations when there is difference of gene frequency among the parents for all loci of the polygenic system, due to the possibility of fixation of favorable genes in all of them.

Whatever the gene distribution in the parents, it is demonstrated that:

\[
\tilde{F}_1 = m + [d] + [i] + [i']
\]

\[
\tilde{F}_1 = m - [d] + [i] - [i']
\]

\[
F = \frac{1}{2} \left[ l \right] + \frac{1}{2} \left[ h \right] + \frac{1}{2} \left[ i \right] + \frac{1}{2} \left[ j \right]
\]

where:

- \(m\) is the genotypic mean of the \(F_0\) generation and, therefore, a parameter defined in this generation; it is the most informative component of means, determined by genic effects which cannot be modified by selection;
- \([d] = \sum_{i=1}^{k_1} d_i + \sum_{i=1}^{k_2} d_i\) is a parameter defined in relation to the parent with greatest phenotypic expression (\(P_1\)); it is determined by the sum of the deviations between the genotypic value of the homozygote with greatest expression and the mean of the homozygotes, in relation to genes of \(P_1\) that are not fixed in the polygenic system under study (additive component); of course, \(k_1\) is the numbers of non fixed genes that increase and decrease trait expression, respectively \((k_1 + k_1 = k_2)\);
- \([i] = \sum_{i=1}^{k_3} i_n + \sum_{i=1}^{k_4} i_n\) is a parameter also defined in relation to \(P_1\), being determined by the sum of the epistatic effects between homozygous combinations of non fixed genes (additive by additive epistatic component);
[i'] = \sum_{k=1}^{r} \sum_{s=1}^{s} l_i - \sum_{k=1}^{r} \sum_{s=1}^{s} l_i is also defined in relation to P_1, being determined by the sum of the epistatic effects between homozygous combination of non fixed gene and homozygous combination of fixed gene (additive by additive epistatic component); [i'] = \sum_{k=1}^{r} \sum_{s=1}^{s} l_i is a parameter defined in relation to the F_1 generation, determined by the sum of the epistatic effects between fixed genic combination and heterozygous genic combination (additive by dominant epistatic component); [l] = \sum_{k=1}^{r} \sum_{s=1}^{s} l_i is a parameter also defined in relation to the F_1 generation, being determined by the sum of the epistatic effects between heterozygous genic combinations (dominant by dominant epistatic component).

As for parents and generations obtained from selfing, the genotypic means of other generations can be expressed in terms of the genetic components m, [d], [h], [l], [i'], [j] and [l], regardless of the gene distribution in the parents. The component [l] is a parameter determined by the sum of the epistatic effects between homozygous combination of non fixed gene and heterozygous genic combination (additive by dominant epistatic component). The genotypic means of the generations obtained by backcrossing are:

\[ B_i = F_s P_x F_i = m + \frac{1}{2} [d] + \frac{1}{4} [h] + \frac{1}{4} [i'] + \frac{1}{4} [j] + \frac{1}{4} [l] \]
\[ B_i = F_s P_x F_i = m - \frac{1}{2} [d] + \frac{1}{4} [h] - \frac{1}{4} [i'] - \frac{1}{4} [j] + \frac{1}{4} [l] \]
\[ B_i = B_s P_x F_i = m + \frac{1}{4} [d] + \frac{1}{4} [h] + \frac{1}{4} [i'] + \frac{1}{4} [j] + \frac{1}{4} [l] \]
\[ B_i = B_s P_x F_i = m + \frac{1}{4} [d] + \frac{1}{4} [h] + \frac{1}{4} [i'] - \frac{1}{4} [j] + \frac{1}{4} [l] \]
\[ B_i = B_s P_x F_i = m + \frac{1}{4} [d] + \frac{1}{4} [h] + \frac{1}{4} [i'] - \frac{1}{4} [j] - \frac{1}{4} [l] \]
\[ B_i = B_s P_x F_i = m + \frac{1}{4} [d] + \frac{1}{4} [h] + \frac{1}{4} [i'] - \frac{1}{4} [j] + \frac{1}{4} [l] \]
\[ B_i = B_s P_x F_i = m + \frac{1}{4} [d] + \frac{1}{4} [h] + \frac{1}{4} [i'] - \frac{1}{4} [j] - \frac{1}{4} [l] \]

The genotypic means of the generations obtained through selfing individuals in B_1 and B_2 are:

\[ B_i = m + \frac{1}{2} [d] + \frac{1}{4} [h] + \frac{1}{4} [i'] + \frac{1}{8} [j] + \frac{1}{16} [l] \]
\[ B_i = m - \frac{1}{2} [d] + \frac{1}{4} [h] + \frac{1}{4} [i'] - \frac{1}{8} [j] + \frac{1}{16} [l] \]

It is further shown that:
\[ F_1 P_x F_i = B_i \]
\[ F_1 P_x F_i = B_2 \]
\[ F_1 P_x F_i = F_2 \]
The genotypic mean of the generation obtained by random mating \( F_2 \) individuals is:

\[ \mathbb{F}_{3_{\text{re}}} = \mathbb{F} \]

And further:

\[ B_{xF_2} = m + \frac{1}{4}d + \frac{1}{2}h + \frac{1}{16}i + \frac{1}{4}j + \frac{1}{8}k + \frac{1}{2}l \]

\[ B_{xF_2} = m - \frac{1}{4}d + \frac{1}{2}h + \frac{1}{16}i - \frac{1}{4}j - \frac{1}{8}k + \frac{1}{2}l \]

The genotypic means of the populations obtained by random mating among individuals of the \( B_1 \) and among individuals of the \( B_2 \) generations are:

\[ \mathbb{F}_{3_{\text{re}}} = m + \frac{1}{2}d + \frac{3}{8}h + \frac{1}{4}i + \frac{3}{16}j + \frac{9}{64}l \]

\[ \mathbb{F}_{2_{\text{re}}} = m - \frac{1}{2}d + \frac{3}{8}h + \frac{1}{4}i - \frac{3}{16}j - \frac{9}{64}l \]

**DISCUSSION**

Although the epistatic effects due to interaction between genic combinations of fixed and non fixed genes contribute to determine the genotypic values of individuals and the mean of any generation or population, the relative importance of the corresponding epistatic components \([i']\) and \([j']\) cannot be assessed, because they are not estimable. The coefficient of the component \([i']\) is the same as the additive component \([d]\) and the coefficient of component \([j']\) is the same as the dominant component \([h]\), in any generation. Clearly, it is possible to estimate the components \((d + [i'])\), \((h + [j'])\), \([i]\), \([j]\), and \([l]\). If there are no fixed genes in the polygenic system under study or if there are no epistatic effects responsible for the \([i']\) and \([j']\) components, the genotypic means of all generations are those presented by Mather & Jinks (1974). However, if \([i']\) and \([j']\) are different from zero, the estimates of the parameters \([d]\) and \([h]\), of a biometrical model that involves only interaction between genic combinations of two loci, will be:

\[ [d] = \sum_{r=1}^{k_1} \left( d_r + \sum_{s=1}^{k_2} \sum_{t=1}^{k_3} d_{rst} \right) \]

\[ [h] = \sum_{r=1}^{k_1} \left( h_r + \sum_{s=1}^{k_2} j_{rs} \right) \]

Therefore, the existence of fixed genes and epistasis in the polygenic system under study may alter the nature of the additive and dominance components that are estimated, resulting in biased estimates even with genetic analysis following the additive-dominant model with epistasis. In the following
paragraphs it will be discussed the magnitudes of the bias in the estimates of \([d]\) and \([h]\), considering the possible types of epistasis. When interaction between genes of two loci is considered, the following kinds of epistasis can occur (Mather & Jinks, 1974; Kearsey & Pooni, 1996):

- **complementary epistasis** (9 : 7): \(d_a = d_b = h_a = h_b = i_{ab} = j_{ab} = j_{ba} = l_{ab}\) (positive dominance) or \(d_a = d_b = -h_a = -h_b = -i_{ab} = j_{ab} = j_{ba} = -l_{ab}\) (negative dominance);
- **duplicate epistasis** (15 : 1): \(d_a = d_b = h_a = h_b = -i_{ab} = -j_{ab} = -j_{ba} = -l_{ab}\) (positive dominance) or \(d_a = d_b = -h_a = -h_b = i_{ab} = j_{ab} = j_{ba} = l_{ab}\) (negative dominance);
- **recessive epistasis** (9 : 3 : 4): \(d_a = h_a\) and \(d_b = h_b = i_{ab} = j_{ab} = j_{ba} = l_{ab}\) (positive dominance) or \(d_a = -h_a\) and \(d_b = -h_b = -i_{ab} = -j_{ab} = -j_{ba} = -l_{ab}\) (negative dominance);
- **dominant epistasis** (12 : 3 : 1): \(d_a = h_a\) and \(d_b = h_b = -i_{ab} = -j_{ab} = -j_{ba} = -l_{ab}\) (positive dominance) or \(d_a = -h_a\) and \(d_b = -h_b = i_{ab} = j_{ab} = j_{ba} = l_{ab}\) (negative dominance);
- **dominant and recessive epistasis** (13 : 3): \(d_a = -d_b = h_a = -h_b = i_{ab} = j_{ab} = j_{ba} = l_{ab}\) (positive dominance) or \(d_a = -d_b = -h_a = h_b = -i_{ab} = -j_{ab} = -j_{ba} = l_{ab}\) (negative dominance);
- **duplicate genes with cumulative effects** (9 : 6 : 1): \(d_a = d_b = h_a = h_b\) and \(i_{ab} = j_{ab} = j_{ba} = l_{ab}\) (positive dominance) or \(d_a = d_b = -h_a = -h_b\) and \(i_{ab} = -j_{ab} = -j_{ba} = -l_{ab}\) (negative dominance);
- **non epistatic genic interaction** (9 : 3 : 3 : 1): \(d_a = h_a\) and \(d_b = h_b\) and \(i_{ab} = j_{ab} = j_{ba} = l_{ab}\) (positive dominance) or \(d_a = -h_a\) and \(d_b = -h_b\) and \(i_{ab} = -j_{ab} = -j_{ba} = -l_{ab}\) (negative dominance).

Aiming to assess the bias in the estimates of the additive and dominance components, due to additive x additive and additive x dominant epistatic effects, involving fixed and non fixed genes, unidirectional dominance and the same type of epistasis for all pairs of genes will be assumed. The exception is the dominant and recessive epistasis. In relation to any three of the \(k\) genes of a polygenic system, it is only possible to define this class of epistasis for two pairs of genes. For the third pair the epistasis is complementary or duplicate. In the following expressions, \(p\) is the coefficient of the relationship between deviation \(d\) of a non fixed gene and the corresponding effect \(i'\) \((i' = pd)\) and between dominance effect and the corresponding effect \(j'\) \((j' = ph)\). The values \(x\) and \(y\) are the proportions of fixed genes that increase and decrease trait expression, respectively \((x = k_2/k_2, y = k_2/k_2, x + y = 1)\).

In the cases of complementary epistasis \((p = 1)\), recessive epistasis, duplicate genes with cumulative effects and non epistatic genic interaction, all with positive dominance, it can be demonstrated that:

\[
\frac{i'}{[d]} = p(k_2, k_2) \quad \text{and} \quad \frac{i'}{[d][i']} = \frac{p(x-y)}{(1/k_2)p(x-y)} \times 100\%
\]

\[
\frac{j'}{[h]} = p(k_2, k_2) \quad \text{and} \quad \frac{j'}{[h][j']} = \frac{p(x-y)}{(1/k_2)p(x-y)} \times 100\%
\]

When it is considered complementary \((p = 1)\) or recessive epistasis, both with negative dominance, the relationships are:
The following relations are true in the cases of duplicate (p = 1) and dominant epistasis, both with negative dominance:

\[
\frac{[i']}{[d]} = -p(k_{2y} - k_{2x}) \quad \text{and} \quad \frac{[i']}{[d+i']} = -\frac{p(x-y)}{(-1/k_y)p(x-y)} \quad 100\%
\]

\[
\frac{[j']}{[h]} = -p(k_{2y} - k_{2x}) \quad \text{and} \quad \frac{[j']}{[h+j']} = -\frac{p(x-y)}{(-1/k_y)p(x-y)} \quad 100\%
\]

Finally, when there is duplicate epistasis (p = 1), dominant epistasis, duplicate genes with cumulative effects or non epistatic genic interaction, all with negative dominance, the expressions are:

\[
\frac{[i']}{[d]} = p(k_{2y} - k_{2x}) \quad \text{and} \quad \frac{[i']}{[d+i']} = \frac{p(x-y)}{(1/k_y)p(x-y)} \quad 100\%
\]

\[
\frac{[j']}{[h]} = p(k_{2y} - k_{2x}) \quad \text{and} \quad \frac{[j']}{[h+j']} = \frac{p(x-y)}{(1/k_y)p(x-y)} \quad 100\%
\]

In relation to a couple of genes, if there is dominant and recessive epistasis (p = 1) with positive dominance, the values of \([i'] / [d], [i'] / ([d] + [i']), [j'] / [h] \) and \([j'] / ([h] + [j']) \) are those presented for complementary epistasis with positive dominance, if the fixed gene decreases trait expression, or complementary epistasis with negative dominance, if the fixed gene increases trait expression. With negative dominance the values are those presented for duplicate epistasis with positive dominance, if the fixed gene increases trait expression, or duplicate epistasis with negative dominance, if the fixed gene decreases trait expression.

In general, it can be inferred that the absolute values of the relations \([i'] / [d] \) and \([j'] / [h] \) are proportional to p and the difference between the numbers of fixed genes that increase and decrease trait expression. Then, assuming p \leq 1, the bias in the estimates of the additive and due to dominance components tends to be greater when the predominant epistasis is complementary or duplicate or dominant and recessive. If p is greater than one the bias will be superior if there is recessive or dominant epistasis, duplicate genes with cumulative effects or non epistatic genic interaction. Independently of the type of prevailing epistasis and direction of the deviations due to dominance, if k_{2y} = k_{2x} (x = y) there is no bias in the estimates of the additive and dominance components, since in this case \([i'] = [j'] = 0 \). When k_{2y} and the absolute value of p(k_{2y} - k_{2x}) are high (hundreds to thousands), the bias in the estimates of the additive and dominance components will be of greater magnitude, because \([d] + [i'] \approx [i'] \) and \([h] + [j'] \approx [j'] \), attaining maximum value.
when $k_2$ or $k_2$ equal to $k_2$, regardless the kind of epistasis and signal of deviations due to dominance. Therefore, the estimated values of $[d]$ and $[h]$ can be approximately equal to the estimates of $[i']$ and $[j']$, respectively, negatively affecting inferences on the inheritance of the quantitative trait under analysis.

If the epistatic effects are of reduced magnitude, comparatively to $d$ and $h$ ($p$ close to zero), the bias in the estimates of $[d]$ and $[h]$ tends to be small, even when the difference $k_2 - k_2$ is not. For example, if $p = 0.001$ and $|k_2 - k_2| = 100$, the absolute values of $[i']$ and $[j']$ correspond to 10% of the absolute values of $[d]$ and $[h]$, respectively. Even when the value of $p(k_2 - k_2)$ is reduced, 1 or -1 for example, the bias in the estimates of the additive and dominance components can be high, 100% ($[d] + [i'] = 2[d]$ and $[h] + [j'] = 2[h]$) or -100% ($[d] + [i'] = 0$ and $[h] + [j'] = 0$), the sign being dependent of the predominant epistasis and the direction of dominance effects.

Lastly, an alternative to avoid biased estimates of $[d]$ and $[h]$ is to use contrasting parents, that is, with means sharply different.

CONCLUSIONS

1. Epistatic effects involving genic combinations of fixed and non fixed genes contribute to the genotypic mean of any population.
2. These effects define specific additive x additive and additive x dominant epistatic components of means.
3. As such components are not estimable, their relative importance cannot be assessed.
4. These epistatic effects can cause bias in the estimates of the additive and dominance components to which they are confounded.
5. The magnitude of the bias depends on the relative values of the epistatic effects, comparatively to deviations $d$ and $h$, type of prevailing epistasis and direction of dominance.

REFERENCES


