Theor. Appl. Genet. 62, 5–7 (1982)

A Non-linear Genetic Model

J. R. Sedcole

Grasslands Division, Department of Scientific and Industrial Research, Palmerston North (New Zealand)

Summary. A model, developed by Seyffert and Forkmann (1976), simulates quantitative characters by genes with biochemically definable action. This model, however, possesses a number of shortcomings which have been overcome by a modified model of the form:

$$y_{[x_1, x_2, \ldots x_k]} = Y^k \prod_{i=1}^{k} [1 - (1 - r_i) \exp(-c_i x_i)]$$

where $y_{[x_1, x_2, \ldots x_k]}$ is the score of the genotype $[x_1, \ldots x_k]$, x_i is the number of positive alleles (0,1,2) at locus i, and Y, r_i, c_i are fitted constants. As well as having a better fit to the data published by Seyffert and Forkmann for the anthocyanin content of flowers of *Matthiola incana*, this modified model has implications concerning heterosis, multiple allelism and optimum genotypes.

Key words: Biochemical action model – Anthocyanin – Heterosis – Multiple allelism – Optimum genotypes

Introduction

Many quantitative genetics models are based on linear expressions which bear little, if any, relation to the biochemical action of the gene or genes the model is supposed to be describing. Seyffert, Forkmann and co-workers (for references see Forkmann and Seyffert 1977), however, have sought to develop models that represent more closely the biochemical action. In one study, Seyffert and Forkmann (1976) related the anthocyanin content of flowers of *Matthiola incana* R. Br. to the number of functional alleles controlling that character by the Baule-Mitscherlich function:

$$\frac{dy}{dx} = c\,(Y - y).$$

This equation was integrated to give:

$$y = Y\,(1 - e^{-c\,x}) \qquad (1)$$

where y is the measured value of the genotype, Y and c are constants, and x is the number of functional alleles controlling the character. They argued that their observations on isogenotypes of *M. incana* implied a saturation curve, and that such curves often characterize enzymatic reactions.

However, the biochemical connection between enzyme kinetics and the phenotype may not always be as direct as Seyffert and Forkmann seem to imply, but the observed relationship between the phenotype and allele number may often be adequately described by the Baule-Mitscherlich function.

Model (1), however, does not distinguish between the effect of alleles at different loci. To overcome this problem Seyffert and Forkmann (1976) suggested a different model:

$$y = Y\,(1 - e^{-c_1 x})\,(1 - e^{-c_2 x}) \ldots (1 - e^{-c_k x}), \qquad (2)$$

but they could not find an appropriate estimation procedure to fit this model to their data. Seyffert and Forkmann (1976) pointed out that (2) is appropriate only for complementary gene action, i.e., even if only one locus has no positive alleles, y is zero, a result at variance with their own data.

The Model

Equation (2) may be modified thus:

$$y_{[x_1, x_2, \ldots x_k]} = Y^k \prod_{i=1}^{k} [1 - (1 - r_i) \exp(-c_i x_i)] \qquad (3)$$

where x_i is the number of positive (rather than "functional") alleles at locus i ($x_i = 0,1,2$) and Y, r_i and c_i are fitted constants. This model preserves the multiplicative nature of (2), it distinguishes between loci, and has the biochemically reasonable saturation curve of (1).

For any locus, if $x_i = 0$, the expression in the brackets becomes $[r_i]$, so that the multiplicative contribution of

0040-5752/82/0062/0005/$ 1.00

Table 1. Observed values (Seyffert and Forkmann 1976) of anthocyanin content of flowers of *Matthiola incana* and values estimated by model (4)

Genotype	Observed values	Estimated values
0,0,0	680	830
0,0,1	743	867
0,1,0	1017	970
0,1,1	1038	1013
1,0,0	1063	966
1,0,1	1006	1008
1,1,0	1122	1129
1,1,1	1199	1179
0,0,2	769	874
0,1,2	1071	1021
1,0,2	983	1016
1,1,2	1154	1188
0,2,0	1059	1000
0,2,1	1125	1044
1,2,0	1121	1163
1,2,1	1210	1215
0,2,2	1108	1052
1,2,2	1242	1224
2,0,0	1203	1018
2,0,1	1112	1063
2,1,0	1191	1190
2,1,1	1218	1243
2,0,2	1096	1072
2,1,2	1199	1253
2,2,0	1058	1227
2,2,1	1251	1281
2,2,2	1350	1291

the homozygous recessive locus i is (Y) $(r_i) = y_{0i}$, say, i.e., $r_i = y_{0i}/Y$. Thus, e.g.,

$$y_{[0,0,0]} = y_{01} \cdot y_{02} \cdot y_{03}.$$

While no analytical solution is available for a least-squares estimation for (3), the modified Gauss-Newton method (Hartley 1961) was found to be a satisfactory numerical approach for finding a least-squares solution for (3) for the data presented by Seyffert and Forkmann (1976). Thus, with such a solution, (3) may be written:

$$y_{[x_1, x_2, x_3]} = 11.04^3 [1 - 0.211\, e^{-0.948 x_1}] \times [1 - 0.276\, e^{-1.553 x_2}] [1 - 0.052\, e^{-1.687 x_3}] \quad (4)$$

Values estimated by (4) are compared with the data given by Seyffert and Forkmann (1976) in Table 1. The residual mean square is 7746, which may be compared with 10,427 from fitting (1).

Discussion

We may first note that the heterozygote, [1,1,1] could arise from any of the crosses between homozygous genotypes [0,0,0] × [2,2,2], [0,0,2] × [2,2,0], [0,2,0] × [2,0,2] or [0,2,2] × [2,0,0]. The data of Seyffert and Forkmann (1976) show that for some of these instances, e.g., [0,2,0] × [2,0,2], the heterozygote outperforms both homozygous "parents" (Table 1). Thus a form of heterosis is observed. The model (1) does not account for this result, whereas, as the tabulated values show, the model (4) does show this heterotic effect. As both the observed data and the model (4) indicate, "heterosis" may not occur for some crosses, even though the heterozygote is the same.

This feature of the modified model emphasises the possibly misleading nature of the linear biometrical model of gene action. If α_i is the additive effect of the ith allele, β_{ij} the intralocus interaction between alleles i, j, then the genotypic values of the homozygous genotypes P_1, P_2 and their F_1 offspring are,

$$\bar{P}_1 = \mu + 2\,\alpha_1$$

$$\bar{P}_2 = \mu + 2\,\alpha_2$$

$$\bar{F}_1 = \mu + \alpha_1 + \alpha_2 + \beta_{12}.$$

Thus the intralocus interaction, β_{12} is implicated as the source of heterosis.

A second point is that the best genotype is that one homozygous for all favourable alleles. This is shown by both the observed data and the model.

Thirdly, multiple allelism can be accommodated by this model. We may take, as a basis, the fact that multiple alleles rarely are mutations that replace each other at a point on the chromosome, but are, rather, mutations at slightly different positions, either within the cistron, or within the complex locus. The model (3) may be expanded so that one bracket represents the effect of one allele within a multiple allelic series. Even if the alleles are point replacements, the model (3) will still satisfactorily characterize the system.

For example, the allelic series a_0, a_1, a_2 may form the genotype $a_0\, a_1$ with the value:

$$y_{[1,1,0]} = Y^3 [1 - (1 - r_0) \exp(-c_0)] \times [1 - (1 - r_1) \exp(-c_1)] [r_2].$$

If an allele completely blocks biochemical action (complementary gene action), then $r_i = 0$, and c_i is some fitted large number for the locus of the allele.

One marked characteristic of the data not accommodated by the modified model is the impairment induced by the addition of the third positive allele at the third locus, e.g., cf. [1,0,1] with [1,0,2]. None of these aberrations are significant (mean S.D. of the observed means is 15), but [2,2,0] is significantly inferior ($P < 0.001$) to [1,2,0] and [2,1,0]. A possible explanation for this aberration is that the genotypes observed were not absolutely isogenic and that other unidentified loci are also involved.

Acknowledgement

I thank Dr Ed Pollak, Iowa State University, for bringing the work of Seyffert and co-workers to my attention.

Literature

Forkmann, G.; Seyffert, W. (1977): Simulation of quantitative characters by genes with biochemically definable action. 6. Modification of a simple model. Genetics **85,** 557–572

Hartley, H.O. (1961): The modified Gauss-Newton method for the fitting of non-linear regression functions by least squares. Technometrics **3,** 267–280

Seyffert, S.; Forkmann, G. (1976): Simulation of quantitative characters by genes with biochemically definable action. 7. Observations and discussion of non-linear relationships. In: Population Genetics and Ecology. (eds. Karlin, S.; Nevo, E.), pp. 431–440. New York: Acad. Press

Received November 20, 1981
Communicated by R. W. Allard

Dr. J. Richard Sedcole
Grasslands Division
Department of Scientific and Industrial Research
Private Bag
Palmerston North (New Zealand)