Un autre regard sur le modèle multiplicatif en génétique quantitative

Christine Dillmann, Jean Louis J. L. Foulley

Original article

Another look at multiplicative models in quantitative genetics

Christine Dillmann [a*], Jean-Louis Foulley [b]

[a] Station de génétique végétale, INA-PG, ferme de Moulon,
91190 Gif-sur-Yvette, France
[b] Station de génétique quantitative et appliquée, Institut national
de la recherche agronomique, 78352 Jouy-en-Josas cedex, France

Abstract – This paper reviews basic theory and features of the multiplicative model of gene action. A formal decomposition of the mean and of the genotypic variance is presented. Connections between the statistical parameters of this model and those of the factorial decomposition into additive, dominance and epistatic effects are also emphasized. General formulae for the genotypic covariance among inbred relatives are given in the case of linkage equilibrium. It is shown that neglecting the epistatic components of variation makes the multiplicative model a pseudo-additive one, since this approximation does not break the strong dependency between mean and variance effects. Similarities and differences between the classical polygenic 'additive-dominance' and the multiplicative gene action approaches are outlined and discussed. Numerical examples for the biallelic case are produced to illustrate that comparison. © Inra/Elsevier, Paris

multiplicative gene action / covariance among relatives / inbreeding

Résumé – Un autre regard sur le modèle multiplicatif en génétique quantitative. Cet article présente la théorie et les principales caractéristiques du modèle multiplicatif d'action des gènes. Une décomposition formelle de la moyenne et de la variance génotypique permet d'établir les relations entre les paramètres statistiques de ce modèle et ceux issus de la décomposition factorielle de l'effet des gènes en effets additifs, de dominance et d'épistasie. Une formule générale de la covariance entre apparentés dans une population consanguine en équilibre de liaison est proposée. On montre que les composantes épistatiques de la variabilité génotypique peuvent être négligées ; le modèle multiplicatif devient alors un modèle pseudo-additif, l'approximation ne supprimant pas la forte liaison entre moyenne et variance. Les similitudes et les différences entre le modèle polygénique « additif-dominance » classique et le modèle multiplicatif d'action des gènes sont discutées et illustrées par des exemples dans le cas biallélique. © Inra/Elsevier, Paris

modèle multiplicatif / covariance entre apparentés / consanguinité

* Correspondence and reprints
E-mail: dillmann@moulon.inra.fr

1. INTRODUCTION

Most models for quantitative characters in evolutionary genetics proceed from a few concepts developed by Fisher [15] in applying Mendel's laws to complex characters: genetic variation is due to a very large number of independent loci whose effects are very small and of about the same magnitude at each locus. The use of a statistical linear decomposition of the genotypic value into mean effects of genes and allelic interaction effects within and between loci justifies the use of a multifactorial model. Furthermore, the assumption of an infinite number of loci without epistasis leads to normal distribution theory with properties that allow prediction of changes in moments of traits for populations subjected to different evolutionary forces such as drift or selection [2]. However, phenotypes may be controlled by other mechanisms of gene action. The optimum model [45], the multiplicative model [24] and the synergistic model [27] have been proposed as alternatives to the additive or additive-dominance models.

One of the basic features of the multiplicative model is that it creates allelic interactions between loci and introduces a dependency between mean and variance of a trait. Indirect evidence for multiplicative gene action (MGA) has been provided by studying the distributions of breeding values for complex traits. Such distributions are expected to be skewed under MGA [35], or to become Gaussian after logarithmic transformation. Such evidence has been found for height [11] and growth [10] in mice and on fruit weight in tomatos [14, 35]. More generally, production traits are multiplicative [21]. Grain yield in maize is the product of the number of seeds and mean weight of seeds. In the same way, prolificacy of domestic mammals is the product of ovulation rate and embryo survival. There is also some experimental evidence for the applicability of such a model to the biosynthesis chain of anthocyanin in flowers [38]. Furthermore, the multiplicative model turns out to be the common choice in resource allocation models, when considering a trade-off between life history traits such as seed and pollen production, or survival rate (biomass) and reproduction [12, 13, 41]).

Theory of the multiplicative model was worked out by Cockerham [7] who proposed a partition of the genotypic variance and expressed the amount of non-additive variation due to multiplicative effects of genes. As he pointed out, non-additive variation was rather small for realistic values of the total genotypic coefficient of variation (less than 40 %). It was suggested that the multiplicative model was formally an additive one. However, this model has been used as an explanation for heterosis [21, 37]. The authors emphasized the analogy for the decomposition of the mean between a multiplicative model at the trait level (one trait being the product of several traits) and a multiplicative model at the gene level (multiplicative gene action). Using generation means, a test for the existence of multiplicative effects was proposed [33] and some evidence for such multiplicative effects was found in fava beans. Finally, considering a trait governed by genes with multiplicative effects and undergoing stabilizing selection, Gimelfarb [19] showed that MGA enhances some 'hidden' variability which can be expressed as additive variation when selection is relaxed. This appears to be a general consequence of epistasis, which has the main effect of modifying the additive and dominance components of the genetic variance

[4]. Another consequence of epistasis is the increase of the additive variation in finite [9, 20] or subdivided [44] populations. In this context, it seems worthwhile to reconsider some implications of the multiplicative model.

The purpose of this paper is threefold: i) to present a formal decomposition of the mean and of the genotypic variance under MGA, ii) to make connections between the statistical parameters of this model and those of the classical decomposition of the genotypic value into its additive, dominance and epistatic components, and iii) to derive exact and approximate formulae for the covariance among inbred relatives under MGA.

2. DECOMPOSITION OF THE GENOTYPIC VALUE

2.1. Classical theory

The decomposition of the genotypic value of an individual was derived by Fisher [15] based on the 'factorial' method of experimentation and later generalized by Kempthorne [25]. It proceeds from the factorial decomposition of genotypic values in a panmictic population of infinite size.

Consider a character determined by S autosomal loci. Let L be the set of all possible genotypes at the S loci, and G_z be the random variable designating the genotypic value of an individual chosen at random in the population with $z \in L$. The realized value g_z can be partitioned into different effects and interactions within and between loci:

$$
\begin{aligned}
g_z = \mu &+ \sum_s (\alpha_{i_s} + \alpha_{j_s}) \\
&+ \sum_s \beta_{ij_s} \\
&+ \sum_s \sum_{t>s} \left[(\alpha\alpha)_{i_s k_t} + (\alpha\alpha)_{i_s l_t} + (\alpha\alpha)_{j_s k_t} + (\alpha\alpha)_{j_s l_t} \right] \\
&+ \dots
\end{aligned}
\tag{1}
$$

where i and k refer to the paternal allelic forms at loci s and t, respectively, and j and l to the maternal allelic forms at loci s and t, respectively; μ is the general mean; α_{i_s} is the average (or additive) effect of allele i at locus s; β_{ij_s} is the first order interaction (or dominance) effect between alleles i and j at locus s; $(\alpha\alpha)_{i_s k_t}$ is the first order interaction (or additive by additive effect) between the additive effects of allele i at locus s and of allele k at locus t.

In a large panmictic population, supposing that all the loci are in linkage equilibrium, the corresponding components of variance are:

$$
\sigma_A^2 = 2 \sum_{s=1}^{S} \left(\sum_i p_{i_s} \alpha_{i_s}^2 \right)
\tag{2a}
$$

$$
\sigma_D^2 = 2 \sum_{s=1}^{S} \left(\sum_i \sum_j p_{i_s} p_{j_s} \beta_{ij_s}^2 \right)
\tag{2b}
$$

$$
\sigma_{AA}^2 = 2 \sum_{s=1}^{S} \sum_{t \neq s} \left(\sum_i \sum_k p_{i_s} p_{k_t} (\alpha\alpha)_{i_s k_t}^2 \right)
\tag{2c}
$$

where σ_A^2, σ_D^2 and σ_{AA}^2 represent the additive, dominance and additive by additive components of variance, respectively, and p_{i_s} is the frequency of allele i at locus s in the population. Other components of the genetic variance, such as the additive by dominance (σ_{AD}^2) and the dominance by dominance (σ_{DD}^2) epistatic variances may be derived in the same way. If the loci are in linkage disequilibrium in the population, extra covariance terms among effects at those loci must be added, and the expression of variance components becomes somewhat complex, especially under selection and assortative mating [2, 28, 40].

2.2. Partition of the mean and variance under MGA

Let A_s be the effect of alleles at locus s for a randomly chosen individual having z as genotype, and a_{ij_s} be the realized value of this random variable given $z = (ij)$ at locus s. Then, by definition of MGA, the genotypic value is

$$G_z = \prod_{s=1}^{S} A_s \tag{3}$$

One can express the mean μ and the variance σ_G^2 of G_z as functions of the mean $\bar{a}_s = \sum_{ij} p_{ij_s} a_{ij_s}$ and variance $\sigma_s^2 = \sum_{ij} p_{ij_s} (a_{ij_s} - \bar{a}_s)^2$ of the A_ss. Under linkage equilibrium, the A_ss are independent so that

$$\mu = \prod_{s=1}^{S} \bar{a}_s \tag{4}$$

Under the same assumption, $E(G_z^2) = \prod_s E(A_s^2)$, and the expression for the variance is

$$\sigma_G^2 = \prod_{s=1}^{S} (\sigma_s^2 + \bar{a}_s^2) - \prod_{s=1}^{S} \bar{a}_s^2 \tag{5a}$$

In equation (5a), σ_G^2 is a product of sums, but may be alternatively expressed as a sum of products of means and variances because the product of mean effects over the S loci cancels out due to statistical independence. Denote by Λ the set of the S loci and Γ the set of all possible subsets of Λ, the null set excepted, and then

$$\sigma_G^2 = \sum_{U \in \Gamma} \left[\left(\prod_{s \in U} \sigma_s^2 \right) \left(\prod_{t \in \Lambda - U} \bar{a}_t^2 \right) \right] \tag{5b}$$

where U stands for any element of Γ. For example, with two loci, $\Gamma = \langle \langle 1 \rangle, \langle 2 \rangle, \langle 1, 2 \rangle \rangle$, and $\sigma_G^2 = \sigma_1^2 \bar{a}_2^2 + \sigma_2^2 \bar{a}_1^2 + \sigma_1^2 \sigma_2^2$.

2.3. Relationships with parameters of the factorial method

This section deals with the different components of genotypic values under MGA resulting from the application of the factorial method. Mathematical details and derivations are given in Appendix A. They follow straightforwardly from the general approach of Kempthorne [26]. Note that a formal decomposition of this model limited to the mean deviation effects has already been given by Schnell and Cockerham [37] for two loci.

Let $\mu_{ij_s} = E(G_z \mid z \equiv ij_s)$ be the conditional expectation of the genotypic value G_z given the (ordered) genotype z being ij at locus s. The additive effect of allele i at locus s is defined as $\alpha_{i_s} = E_j(\mu_{ij_s}) - \mu$. Using equation (4) for μ and factoring $\prod_{t \neq s} \bar{a}_t$, this effect can be expressed as $\alpha_{i_s} = \left(\sum_j p_{j_s} a_{ij_s} - \bar{a}_s \right) \left(\prod_{t \neq s} \bar{a}_t \right)$.

The first term can be interpreted as an additive effect among the a_{ij_s} values at locus s. Denote this effect by

$$\alpha_{i_s}^* = \sum_j p_{j_s} a_{ij_s} - \bar{a}_s \tag{6a}$$

Then,

$$\alpha_{i_s} = \alpha_{i_s}^* \left(\prod_{t \neq s} \bar{a}_t \right) \tag{6b}$$

Thus, under MGA, the additive effect of an allele at locus s is the product of the additive effect of the allele among the effects of genotypes at locus s times the product of mean genetic effects at the other loci.

Similarly, the dominance effect β_{ij_s} between alleles i and j at locus s is the product of $\prod_{t \neq s} \bar{a}_t$ and the dominance effect $\beta_{ij_s}^*$ among the a_{ij_s} at locus s, i.e.

$$\beta_{ij_s}^* = a_{ij_s} - \alpha_{i_s}^* - \alpha_{j_s}^* - \bar{a}_s \tag{7a}$$

and

$$\beta_{ij_s} = \beta_{ij_s}^* \left(\prod_{t \neq s} \bar{a}_t \right) \tag{7b}$$

Using equations (6a) and (6b), the additive by additive effect $(\alpha\alpha)_{i_s k_t}$ pertaining to allele i at locus s and allele k at locus t is:

$$(\alpha\alpha)_{i_s k_t} = \alpha_{i_s}^* \alpha_{k_t}^* \left(\prod_{u \neq s,t} \bar{a}_u \right) \tag{8}$$

Thus, in the multiplicative model, the genetic components (α, β) at a locus level depend upon the mean genotypic values at other loci. More precisely, any interaction effect can be expressed as the product of the additive and dominance effects among the genotypic effects at each locus times the product of mean genotypic effects at different loci. For instance, the additive by dominance $(\alpha\beta)$, dominance by dominance $(\beta\beta)$ and additive by additive by additive $(\alpha\alpha\alpha)$ epistatic components can be written as:

$$(\alpha\beta)_{i_s kl_t} = \alpha_{i_s}^* \beta_{kl_t}^* \left(\prod_{u \neq s,t} \bar{a}_u \right) \tag{9a}$$

$$(\beta\beta)_{ij_s kl_t} = \beta_{ij_s}^* \beta_{kl_t}^* \left(\prod_{u \neq s,t} \bar{a}_u \right) \tag{9b}$$

$$(\alpha\alpha\alpha)_{i_s k_t m_u} = \alpha_{i_s}^* \alpha_{k_t}^* \alpha_{m_u}^* \left(\prod_{v \neq s,t,u} \bar{a}_v \right) \tag{9c}$$

Using formulae (6b), (7b), (8) and (9abc), one can derive the expression for the different variance components (see Appendix B). The additive genetic variance (σ_A^2) is the sum, over all loci, of the product of the additive genetic variance $(\sigma_{a_s}^2)$ among the a_{ij_s} values at each locus s times the product of the squared mean effects at the other loci:

$$\sigma_A^2 = \sum_{s=1}^{S} \left[\sigma_{a_s}^2 \left(\prod_{t \neq s} \bar{a}_t^2 \right) \right] \tag{10a}$$

where

$$\sigma_{a_s}^2 = 2 \sum_i p_{i_s} \alpha_{i_s}^{*\,2} \tag{10b}$$

Note that equation (10a) can alternatively be written as $\sigma_A^2 = \mu^2 \sum_{s=1}^{S} \left(\dfrac{\sigma_{a_s}^2}{\bar{a}_s^2} \right)$, for $\bar{a}_s \neq 0$. This shows that, under MGA, variance components are related to squared coefficients of variation at each locus.

Similarly, the dominance variance (σ_D^2) can be expressed as the sum, over loci, of the dominance variance $(\sigma_{d_s}^2)$ among the $\beta_{ij_s}^*$ elements times the product of the squared mean effects at the other loci:

$$\sigma_D^2 = \sum_{s=1}^{S} \left[\sigma_{d_s}^2 \left(\prod_{t \neq s} \bar{a}_t^2 \right) \right] \tag{11a}$$

where

$$\sigma_{d_s}^2 = 2 \sum_i \sum_j p_{i_s} p_{j_s} \beta_{ij_s}^{*\,2} \tag{11b}$$

The additive by additive epistatic variance reduces to:

$$\sigma_{AA}^2 = \frac{1}{2} \sum_{s=1}^{S} \sigma_{a_s}^2 \left[\sum_{t \neq s} \sigma_{a_t}^2 \left(\prod_{u \neq s,t} \bar{a}_u^2 \right) \right] \tag{12a}$$

while the additive by dominance (σ_{AD}^2), the dominance by dominance (σ_{DD}^2) and the additive by additive by additive (σ_{AAA}^2) genetic variances are:

$$\sigma_{AD}^2 = \frac{1}{2} \sum_{s=1}^{S} \sigma_{a_s}^2 \left[\sum_{t \neq s} \sigma_{d_t}^2 \left(\prod_{u \neq s,t} \bar{a}_u^2 \right) \right] \tag{12b}$$

$$\sigma_{DD}^2 = \frac{1}{2} \sum_{s=1}^{S} \sigma_{d_s}^2 \left[\sum_{t \neq s} \sigma_{d_t}^2 \left(\prod_{u \neq s,t} \bar{a}_u^2 \right) \right] \tag{12c}$$

$$\sigma_{AAA}^2 = \frac{1}{3!} \sum_{s=1}^{S} \sigma_{a_s}^2 \left[\sum_{t \neq s} \sigma_{a_t}^2 \left[\sum_{u \neq s,t} \sigma_{a_u}^2 \left(\prod_{v \neq s,t,u} \bar{a}_v^2 \right) \right] \right] \tag{12d}$$

Hence, each variance component can be easily expressed as the combination of a genetic variance at one locus (or product of variances at different loci) times squared mean effects at the remaining loci. The total genetic variance as defined in equations (5a) and (5b) can be decomposed as the sum of all

such partitions. The highest order variance corresponds to the $(S-1)$th order interaction

$$\sigma^2_{A^{S-1}} = \prod_{s=1}^{S} \sigma^2_{a_s} \tag{13}$$

Table I illustrates the partition of the genetic variance as expressed analytically in formulae (10ab), (11ab), (12abc) for a trait controlled by MGA. Clearly, the additive and dominance components of variance depend not only on additive and dominance genetic effects at each locus but also on average genotypic values at the other loci.

2.4. Covariances between arbitrary relatives

Extensions of those formulae to covariances between relatives can be easily derived. De Jong and Van Noordwijk [12] gave the expressions for covariances between non-inbred relatives and between life-history traits for some models of resource allocation. Those results are now extended to the case of inbreeding.

We consider here the variability of a neutral trait governed by independent loci in a large, possibly inbred, population. Under those hypotheses, the loci are expected to be in linkage equilibrium and the A_ss are independent, so that $E(G_z) = \prod_{s=1}^{S} E(A_s)$. Now, $E(A_s) = \bar{a}_{F_s} = \bar{a}_s + f_z d_{0_s}$, where f_z is the probability of identity by descent between two homologous alleles of an individual (denoted here as z) drawn at random in the population, and $d_{0_s} = \sum_i p_{i_s} \beta^*_{ii_s}$ is the average dominance effect in the homozygous population. Therefore, the first moment of the distribution of G_z is

$$\mu = \prod_{s=1}^{S} \left(\bar{a}_s + f_z d_{0_s} \right) \tag{14a}$$

Under the same assumptions and using the same notation as in equation (B.1), the genotypic covariance between two individuals z and z' is defined as:

$$\mathrm{Cov}(G_z, G_{z'}) = \prod_{s=1}^{S} E\left[\left(\bar{a}_s + \alpha^*_{1_s} + \alpha^*_{2_s} + \beta^*_s \right) \left(\bar{a}_s + \alpha^{*'}_{1_s} + \alpha^{*'}_{2_s} + \beta^{*'}_s \right) \right]$$

$$- \prod_{s=1}^{S} \bar{a}_{F_s} \bar{a}_{F'_s} \tag{14b}$$

Hence, the problem reduces to calculating the covariance between relatives at one locus.

Following the basic results obtained by Fisher [15], Wright [45] and Malecot [30], the general expression for covariances between relatives was first derived by Cockerham [5] and Kempthorne [25, 26] under the assumptions of random mating and linkage equilibrium. The case of linked loci was investigated later by Cockerham [6] and Schnell [36]. The case of inbred relatives was independently solved by Harris [22], Gillois [18] and later on by Cockerham [8], assuming the absence of linkage. Using Gillois' identity by descent coefficients, the genotypic

Table I. Partition of the genetic variance under MGA, and comparison with the corresponding parameters of the factorial decomposition (AGA).

Description	Among a_{ij} effects at locus s	Overall loci MGA	AGA
Mean	$\bar{a}_s = \sum_{ij} p_{ij_s} a_{ij_s}$	$\mu = \prod_{s=1}^{S} \bar{a}_s$	$\mu = \sum_{s=1}^{S} \bar{a}_s$
Additive effect	$\alpha^*_{i_s} = \sum_j p_j a_{ij_s} - \bar{a}_s$	$\alpha_{i_s} = \alpha^*_{i_s} \prod_{t \neq s} \bar{a}_t$	$\alpha_{i_s} = \alpha^*_{i_s}$
Dominance effect	$\beta^*_{ij_s} = a_{ij_s} - \alpha^*_{i_s} - \alpha^*_{i_s} - \bar{a}_s$	$\beta_{ij_s} = \beta^*_{ij_s} \prod_{t \neq s} \bar{a}_t$	$\beta_{ij_s} = \beta^*_{ij_s}$
Add. by add. epistatic effect		$\alpha\alpha_{i_s k_t} = \alpha^*_{i_s} \alpha^*_{k_t} \prod_{u \neq t,s} \bar{a}_u$	$\alpha\alpha_{i_s k_t} = \sum_j \sum_l p_j p_l (\mu_{ij_s kl_t}) - \alpha_{i_s} - \alpha_{k_t} - \mu$
Additive variance	$\sigma^2_{a_s} = 2 \sum_i p_i \alpha^{*2}_{i_s}$	$\sigma^2_A = \sum_{s=1}^{S} \sigma^2_{a_s} \prod_{t \neq s} \bar{a}^2_t$	$\sigma^2_A = \sum_{s=1}^{S} \sigma^2_{a_s}$
Dominance variance	$\sigma^2_{d_s} = \sum_i \sum_j p_i p_j \beta^{*2}_{ij_s}$	$\sigma^2_D = \sum_{s=1}^{S} \sigma^2_{d_s} \prod_{t \neq s} \bar{a}^2_t$	$\sigma^2_D = \sum_{s=1}^{S} \sigma^2_{d_s}$
Add. by add. epistatic variance		$\sigma^2_{AA} = \frac{1}{2} \sum_{s=1}^{S} \sigma^2_{a_s} \left[\sum_{t \neq s} \sigma^2_{a_t} \prod_{u \neq t,s} \bar{a}^2_u \right]$	$\sigma^2_{AA} = 4 \sum_{s=1}^{S} \sum_{t \neq s} \sum_i \sum_k p_i p_k (\alpha\alpha)^2_{i_s k_t}$

covariance between two individuals (z, z') from an inbred population under MGA can be written as:

$$\text{Cov}(G_z, G_{z'}) = \prod_{s=1}^{S} \{\bar{a}_{F_s}\bar{a}_{F'_s} + 2\varphi_{zz'}\sigma_{a_s}^2 + (\Delta_9 + \Delta_{12})\sigma_{d_s}^2 + \Delta_1\sigma_{d_{0_s}}^2$$

$$+ \left[\Delta_1 + \frac{1}{4}(\Delta_2 + \Delta_3 + \Delta_4 + \Delta_5)\right]\sigma_{ad_{0_s}} + (\Delta_1 + \Delta_6 - f_z f_{z'})d_{0_s}^2\}$$

$$- \prod_{s=1}^{S} \bar{a}_{F_s}\bar{a}_{F'_s} \tag{15a}$$

where the Δ_is are the probabilities of identity modes; $\varphi_{zz'}$ is Malecot's coefficient of kinship between individuals (z, z'); $\sigma_{a_s}^2 = 2\sum_i p_{i_s}\alpha_{i_s}^{*\,2}$ and $\sigma_{d_s}^2 = 2\sum_{ij} p_{i_s}p_{j_s}\beta_{ij_s}^{*\,2}$ are the classical additive and dominance components of the genetic variance, respectively, in a large panmictic population under linkage equilibrium; $\sigma_{d_{0_s}}^2 = 2\sum_i p_{i_s}\beta_{ii_s}^{*\,2}$ and $d_{0_s}^2$ are the variance and squared mean of the dominance effects, respectively, in the homozygous population, and $\sigma_{ad_{0_s}} = 4\sum_i p_{i_s}\alpha_{i_s}^*\beta_{ii_s}^*$ is the covariance between additive and dominance effects in the same population.

Formula (15a) encompasses three new moment parameters defined in the homozygous population resulting from the condition of full identity between homologous genes. This formula also involves six functions of the elementary identity coefficients. Using for instance identity measures introduced by Zhao-Bang Zeng and Cockerham [46], i.e. $\gamma_1 = \Delta_1 + \frac{1}{4}(\Delta_2 + \Delta_3 + \Delta_4 + \Delta_5)$, $\delta_1 = \Delta_9 + \Delta_{12}$, $\delta_2 = \Delta_1$ and $\delta_3 = \Delta_1 + \Delta_6$, it can be alternatively written in a more condensed form as:

$$\text{Cov}(G_z, G_{z'}) = \prod_{s=1}^{S} \{\bar{a}_{F_s}\bar{a}_{F'_s} + 2\varphi_{zz'}\sigma_{a_s}^2 + \delta_1\sigma_{d_s}^2 + \delta_2\sigma_{d_{0_s}}^2$$

$$+ \gamma_1\sigma_{ad_{0_s}}^2 + (\delta_3 - f_z f_{z'})d_{0_s}^2\} - \prod_{s=1}^{S} \bar{a}_{F_s}\bar{a}_{F'_s} \tag{15b}$$

The same reasoning applies to the genotypic variance and leads to

$$\text{Var}(G_z) = \prod_{s=1}^{S} \{(\bar{a}_{F_s})^2 + (1 + f_z)\sigma_{a_s}^2 + (1 - f_z)\sigma_{d_s}^2$$

$$+ f_z\sigma_{d_{0_s}}^2 + f_z\sigma_{ad_{0_s}} + f_z(1 - f_z)d_{0_s}^2\} - \prod_{s=1}^{S} (\bar{a}_{F_s})^2 \tag{15c}$$

The expressions for variances and covariances between inbred relatives in (15abc) are products of sums and may be decomposed as in (10a–13) into sums of products. This would lead to five components of variance of the first

order, $5S(S-1)$ epistatic components of the second order, $5S(S-1)(S-2)$ epistatic components of the third order and so on. Each variance component is the sum of variances at one or more loci times squared mean effects at the remaining loci. Each covariance component is the sum of covariances at each locus times products of mean effects at the remaining loci.

With our assumptions (independent loci, infinite size population), those expressions are much simpler than the corresponding expressions in the full factorial decomposition [16, 43]. In fact, the only coefficients of identity that are needed in (15bc) are obviously the ones corresponding to identities between four alleles at a single locus.

Under linkage disequilibrium, additional covariances (between loci) occur in the expression of total genetic variance, the effects of which on the variances and covariances are essentially unknown. For instance, the total genetic variance comprises five components for two polymorphic loci in the absence of dominance effects, i.e. an additive variance (σ_A^2), an additive by additive epistatic variance (σ_{AA}^2), covariances between additive effects ($\sigma_{A,A}^2$), between additive and epistatic effects ($\sigma_{A,AA}^2$), and between epistatic effects ($\sigma_{AA,AA}^2$). In the general case, higher order terms are also involved. Note that in this case, the corresponding expression for the population mean (14a) also involves high order identity by descent coefficients.

2.5. Approximations

In this section, we will show that neglecting the epistatic variance components leads to much simpler expressions for the covariance between relatives.

In a panmictic population and under MGA, the ratio of the non-epistatic variance components to the total genotypic variance depends on the number of loci and on the total genotypic coefficient of variation ($CV = \sigma_G/\mu$) [7]. From (5a), $\left(1 + CV^2\right) = \prod_{s=1}^{S}\left(1 + \frac{\sigma_{a_s}^2}{\bar{a}_s^2} + \frac{\sigma_{d_s}^2}{\bar{a}_s^2}\right)$. In the symmetrical case with allelic effects and frequencies being the same at each locus, it reduces to $\left(1 + CV^2\right) = \left(1 + \frac{\sigma_a^2}{\bar{a}^2} + \frac{\sigma_d^2}{\bar{a}^2}\right)^S$ [7]. Using equations (10a) and (11a) and rearranging, we obtain $\dfrac{\left(\sigma_A^2 + \sigma_D^2\right)}{\sigma_G^2} = S\dfrac{\left(1 + CV^2\right)^{1/S} - 1}{CV^2}$. Whatever the number of loci, this ratio is very close to 1 for values of the total genotypic coefficient of variation lower than 40 %. Hence $\sigma_G^2 \approx \sigma_A^2 + \sigma_D^2$ and the total genotypic variance can be approximated by its first order components:

$$\sigma_G^2 \approx \sum_{s=1}^{S}\left[\left(\sigma_{a_s}^2 + \sigma_{d_s}^2\right)\left(\prod_{t\neq s}\bar{a}_t^2\right)\right] \tag{16a}$$

Similar approximations to that given in equation (16a) apply in the case of inbreeding. The covariance between inbred relatives reduces to

$$\text{Cov}(G_z, G_{z'}) \approx \sum_{s=1}^{S} \left[(2\varphi_{zz'}\sigma_{a_s}^2 + \delta_1 \sigma_{d_s}^2 + \delta_2 \sigma_{d_{0_s}}^2 \right.$$
$$\left. + \gamma_1 \sigma_{ad_{0_s}} + (\delta_3 - f_z f_{z'})d_{0_s}^2 \right) \left(\prod_{t \neq s} \bar{a}_{F_t} \bar{a}_{F'_t} \right) \right] \qquad (16b)$$

and the genotypic variance may be approximated by

$$\text{Var}(G_z) \approx \sum_{s=1}^{S} \left[((1+f_z)\sigma_{a_s}^2 + (1-f_z)\sigma_{d_s}^2 + f_z \sigma_{d_{0_s}}^2 \right.$$
$$\left. + f_z \sigma_{ad_{0_s}} + f_z(1-f_z)d_{0_s}^2 \right) \left(\prod_{t \neq s} (\bar{a}_{F_t})^2 \right) \right] \qquad (16c)$$

Note that the approximations (16abc) are tantamount to assuming that the genotypic value G_z can be written (apart from a constant) as

$$G_z \approx \sum_{s=1}^{S} \left[\left(\alpha_{i_s}^* + \alpha_{j_s}^* + \beta_{ij_s}^* \right) \left(\prod_{t \neq s} \bar{a}_{F_t} \right) \right] \qquad (17)$$

Those approximations will be checked numerically in the next section.

Formally, as outlined by Cockerham [7], this approximation makes the multiplicative model an additive one without epistasis and the two could not a priori be distinguished from data. However, it does not break down the dependency between mean and variance, which is one of the main characteristics of the presence of epistasis: the genotypic variance is a sum of products of means and variances at different loci. In other words, the genetic variance at each locus is weighted by mean effects at the other loci. As inbreeding affects both mean and variance effects at each locus (equations 14a and 16bc), it should be possible to distinguish between the two models by comparing different levels of inbreeding for the same population.

3. NUMERICAL RESULTS: THE BIALLELIC CASE

Numerical results presented here rely upon a biallelic symmetrical model. S loci in linkage equilibrium are considered, with allelic frequencies being the same at each locus in the base panmictic population. Genotypic effects of the three possible genotypes at one locus were set to $M+a$, $M+d$ and $M-a$, where M is the mid-parent value and a and d the additive and dominance deviation, respectively. Parameter values for M, a and d were assumed to be the same at each locus. We defined $\delta = d/a$ as the constant degree of dominance.

3.1. Base population

In the base panmictic population, the genetic variance at one locus s is

$$\sigma_s^2 = 2pqa^2 \left[1 - 2(p-q)\delta + (p^2+q^2)\delta^2 \right]$$

and the mean effect of locus s is $\bar{a}_s = M + (p - q)a + 2pqd$ where p is the frequency of the favourable allele and $q = 1 - p$. Under MGA and with allelic effects and allelic frequencies being the same at each locus, the total genetic variance is given by equation (5a) so that

$$\sigma_s^2 = \bar{a}_s^2 \left[\left(1 + CV^2 \right)^{\frac{1}{S}} - 1 \right]$$

Equating these two formulae for σ_s^2 allows us to express the additive deviation (a) as a function of the mean (μ), the total genotypic coefficient of variation (CV), the allelic frequency (p) and the degree of dominance (δ):

$$a = \mu^{\frac{1}{S}} \left[\frac{\left(1 + CV^2 \right)^{\frac{1}{S}} - 1}{2pq \left(1 - 2(p - q)\delta + (p^2 + q^2)\delta^2 \right)} \right]^{\frac{1}{2}}$$

Similarly, M is given by

$$M = \mu^{\frac{1}{S}} - (p - q)a - 2pqd$$

The total genotypic coefficient of variation of the base population is assumed to be equal to $CV_0 = 0.2$, and the mean of the base population is equal to $\mu = 1$. We also took $\delta = 1$ corresponding to complete dominance at each locus, or $\delta = 3$ corresponding to overdominance. Using the approximations in equation (16abc), the total genetic variance of the base population is $\text{Var}(G_0) \approx S\sigma_s^2 \bar{a}_s^{2(S-1)}$, and the initial squared coefficient of variation is $CV_0^2 \approx S \frac{\sigma_s^2}{\bar{a}_s^2}$.

3.2. Inbred population

We consider now an inbred population derived from the base population by changing the reproductive behaviour of the individuals and forcing inbred matings during t generations. In this case, the five variance components in the biallelic model were given by Chevalet and Gillois [3] and Mather and Jinks [32]. They can easily be expressed as functions of σ_s^2, the contribution of one locus to the genetic variance. Let us define h^2, the heritability (narrow sense) as

$$h^2 = \frac{\sigma_{a_s}^2}{\sigma_s^2} = \frac{\left[1 - (p - q)\delta \right]^2}{1 - 2(p - q)\delta + (p^2 + q^2)\delta^2}$$

so that $\sigma_{a_s}^2 = h^2\sigma_s^2$ and $\sigma_{d_s}^2 = d_{0_s}^2 = (1 - h^2)\sigma_s^2$. Similarly, one can define $b = \frac{(p - q)^2}{pq}$ and $r = \frac{(p - q)\delta}{1 - (p - q)\delta}$ so that $\sigma_{d_{0_s}}^2 = b(1 - h^2)\sigma_s^2$ and $\sigma_{ad_{0_s}} = rh^2\sigma_s^2$, respectively.

Therefore, from equation (16abc), the variance of allelic effects at one locus in an inbred population is

$$\sigma_{z_s}^2 = \left[1 + f_z\left(h^2 + b(1 - h^2) + rh^2\right) + f_z^2(h^2 - 1)\right]\sigma_s^2$$

and the covariance of allelic effects at one locus between inbred relatives is

$$\sigma_{zz'_s} = \left[2\varphi_{zz'}h^2 + \delta_1(1 - h^2) + \gamma_1 rh^2 + \delta_2 b + (\delta_3 - f_z f_{z'})(1 - h^2)\right]\sigma_s^2$$

The total genetic variance of the population also depends on

$$m = \frac{\bar{a}_{F_s}}{\bar{a}_s} = 1 + 2f_z\frac{pq(p - q)a\delta}{M + (p - q)a + 2pqa\delta}$$

At any time t, the expected genetic variance is equal to

$$\text{Var}^{\text{MGA}}(G_t) = \left[1 + f_z\left(h^2 + b(1 - h^2) + rh^2\right)\right.$$

$$\left. + f_z^2(h^2 - 1)\right]m^{2(S-1)}\text{Var}(G_0) \qquad (18a)$$

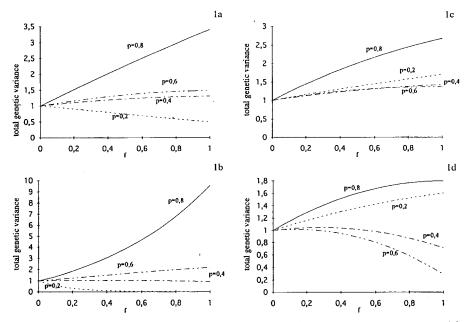while under AGA, $\text{Var}(G_t)$ reduces to [3]:

$$\text{Var}^{\text{AGA}}(G_t) = \left[1 + f_z\left(h^2 + b(1 - h^2) + rh^2\right)\right.$$

$$\left. + f_z^2(h^2 - 1)\right]\text{Var}(G_0) \qquad (18b)$$

Note that the terms between brackets in the right hand sides of equation (18ab) depend only on allelic frequencies and on the degree of dominance. It would therefore be possible to compare the two models of gene action by expressing the genetic variances at time t in units of the total genetic variance in the base population. It is also clear from equation (18ab) that in the biallelic case, the only difference between AGA and MGA is the coefficient $m^{2(S-1)}$ weighting the inbreeding variance under MGA.

Figure 1 illustrates the evolution of the expected total genetic variance with the mean inbreeding coefficient of the populations. Under AGA with complete dominance (*figure 1c*), the genetic variance increases with f_z. The more the frequencies of the favourable alleles depart from 1/2, the higher are the values of the total genetic variance. Under AGA with overdominance (*figure 1d*), the genetic variance decreases with f_z for intermediate frequencies of the favourable allele ($p = 0.4$ or 0.6). Under MGA with complete dominance, two major differences occur as compared to the AGA case. First, for low frequencies of the favourable allele, the genetic variance decreases with inbreeding (*figure 1a*), as seen in the case of AGA with overdominance. Second, the genetic variance increases with the frequency of the favourable alleles. This phenomenon is enhanced by an increase in the number of loci governing the trait (*figure 1b*).

However, such results rely upon the approximations made in equation (16bc), the validity of which is checked by calculating the epistatic components of the
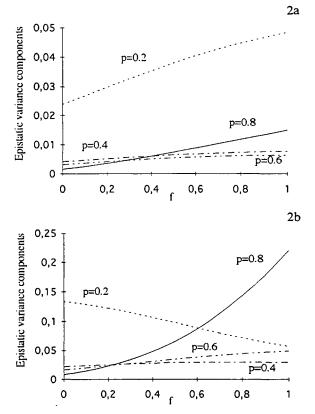
Figure 1. Biallelic case: evolution of the expected total genetic variance with inbreeding for different models of gene action*. a) MGA, 20 loci, $\delta = 1$; b) MGA, 100 loci, $\delta = 1$; c) AGA, 20 loci, $\delta = 1$; d) AGA, 20 loci, $\delta = 3$; * p is the constant frequency of favourable alleles. The total genetic variance is expressed in units of the total genetic variance in the base panmictic population.

total genetic variance at time t, i.e.

$$\text{Var}_{EP}(t) = \frac{\left(\left(\sigma_{z_s}^2 + \bar{a}_{F_s}^2\right)^S - \bar{a}_{F_s}^{2S}\right) - \text{Var}(G_t)}{\text{Var}(G_0)} \text{Var}(G_0) \qquad (19)$$

with $\left(\sigma_{z_s}^2 + \bar{a}_{F_s}^2\right)^S - \bar{a}_{F_s}^{2S}$ corresponding exactly to the definition of the total genetic variance of the population.

Figure 2a and *b* shows the residual epistatic component of the total genetic variance as a function of the mean inbreeding coefficient, f_z, and for the same genetic models as in *figure 1a* and *b*, respectively. It can be seen that the magnitude of the epistatic variance components depends on allelic frequencies. For very low frequencies of the favourable alleles, the epistatic variance never exceeds 10 % of the additive and dominance components. It drops to less than 1 % of the additive and dominance components in other cases and may consequently be neglected.

2a

2b

Figure 2. Biallelic case: evolution of the epistatic components of the genetic variance with inbreeding*. a) MGA, 20 loci, $\delta = 1$; b) MGA, 100 loci, $\delta = 1$; * p is the constant frequency of favourable alleles. The epistatic components of the genetic variance are expressed in units of the total genetic variance in the base panmictic population.

4. DISCUSSION

This study was primarily concerned with the statistical properties of the multiplicative model of gene action. This model can be seen as a good candidate to explain some features of traits observed in physiological or biochemical studies, as well as in classical quantitative genetics experiments (see Introduction).

It was found, as predicted by Cockerham [7], that the epistatic components of variance can be neglected under MGA when the total genotypic coefficient of variation is not too large. It is then possible to describe a trait by invoking only additive and dominance effects at each locus. Tractable formulae for variance components are obtained by neglecting the terms corresponding to products of variances at each locus. Those approximations were shown to be valid even under strong inbreeding. Formally, they make the multiplicative model a pseudo-additive one, since they do not break down the dependency between

mean and variance under MGA. Allelic effects and variances at each locus are weighted by the effect of other loci. This phenomenon may have two main consequences.

First, means and variances at each locus are not affected in the same way by inbreeding. Let us define $CV_A^2 = \sum_s \dfrac{\sigma_{a_s}^2}{\bar{a}_{F_s}^2}$, $CV_{AD_0}^2 = \sum_s \dfrac{\sigma_{ad_{0s}}^2}{\bar{a}_{F_s}^2}$, and so on. It turns out from equation (16b) that the squared total genotypic coefficient of variation of an inbred population can be expressed as

$$\frac{\mathrm{Var}(G_z)}{\mu_z^2} \approx (1+f_z)CV_A^2 + (1-f_z)CV_D^2 + f_z CV_{AD_0}^2 + f_z CV_{D_0}^2 + f_z(1-f_z)CV_{D_0^2}^2$$

This formula clearly indicates that the coefficients affecting the CVs in the right hand side of the above equation are the same as those affecting variance components under AGA. This means that under inbreeding, the formal similarity betwen MGA and AGA is not at the variance level, but rather at the level of the total genotypic coefficient of variation. Such results have been observed in alfalfa by Gallais [17], who pointed out that the genotypic coefficient of variation provided a better scale than the genetic variance to linearize the effect of inbreeding depression on the genetic variation. Note that the analytical expression obtained here relies on some strong assumptions (linkage equilibrium and infinite population) which are discussed below.

Second, if a favourable allele is fixed by selection in a given population, fixation will increase the mean effect of the locus and decrease its variance. These two effects may cancel out in equation (5a) and result in no change for the total genetic variance of the trait. Under AGA, the same phenomenon would lead to a systematic decrease of the total genetic variance.

Despite such important qualitative differences, the two models can hardly be distinguished. In an outbred population, the absence of a significant amount of epistatic variation may be interpreted in two different ways: as originating from polygenic additive-dominance genetic determinism or from multiplicative gene action. Similarly, it is difficult to distinguish between AGA with overdominance and MGA without overdominance in the presence of inbreeding. Note that multiplicative gene action can also be viewed as a parsimonious explanation for heterosis: complete dominance under MGA can explain some patterns of change of the inbreeding genetic variance as does overdominance under AGA.

It is nevertheless possible to test multiplicative gene action by comparing different levels of inbreeding for the same population. Melchinger et al. [33] defined a multiplicative factor and proposed a test based on the comparisons of the means of different inbred generations. Our results suggest a possible test at the variance level restricted to populations exhibiting low frequencies of the favourable alleles. In this case, whatever the degree of dominance, the genetic variance is expected to increase with inbreeding under AGA, and to decrease under MGA.

The numerical results presented here in the biallelic case were obtained by assuming equal allelic effects and frequencies for each locus. The main reason for that was to simplify the complex algebra generated by MGA. However, this assumption should not alter the general trend of the results. We checked numerically that even with strong discrepancies between gene effects, the

epistatic components rarely exceed 10 % of the total genetic variance, as long as the total genotypic coefficient of variation does not exceed 40 % (results not shown). Up to now, the exact distribution of allelic effects over loci governing a trait is not known, even though results concerning the distribution of QTLs, which can be detected in a population, seem to indicate a L-shaped distribution [29, 31, 34]. Relying upon QTL detection results, unequal gene effects may concern a maximum of 20 % of loci that govern a given trait. Equal allelic effect is an implicit assumption in the polygenic additive-dominance model [2, 15]. In our opinion, the strongest assumption here is the equal frequencies hypothesis. Even without random genetic drift, and with the same selection pressure acting on each locus, the allelic frequencies may not be expected to be the same because of mutation.

Most of the results presented here are also heavily dependent on the hypothesis of statistical independence between loci. This hypothesis restricts the analysis to the case of independent loci and large populations. However, such situations may exist in artificial inbred populations created by breeders. In plant breeding, for example, populations of 300 to 500 reproducing individuals are common, with linkage disequilibrium restricted to loci situated on the same chromosome (Dillmann and Charcosset, pers. comm.). In general, random genetic drift in finite populations, as well as linkage between loci, generates multilocus identities by descent [16, 39, 42, 43]. In that case, the validity of our approximations remains to be checked. But, equation (16b) stresses the importance of mean effects at each locus in evolutionary processes, when epistasis is involved, and provides a good basis to study the evolution of genetic variation under inbreeding for MGA traits.

As for the effect of inbreeding, we were only concerned with expectations of the parameters. Those parameters also have a variance which may be calculated [46]. As experimental studies always involve a finite number of populations, and often a unique one, the variation around expected values may be important. In particular, genetic drift and selection generate not only variation between populations in mean performance, but also in within population variance which contributes indirectly to the variation in selection response [1, 23]. It makes the intensity of selection fluctuate and therefore changes the population means at the next generation. Due to the interaction between mean and variance, those fluctuations may even be enhanced by multiplicative gene action. We are presently studying the combined effects of selection and random genetic drift, including also the case of linked loci.

ACKNOWLEDGEMENTS

REFERENCES

[1] Avery P.J., Hill W.G., Variability in genetic parameters among small populations, Genet. Res. Camb 29 (1978) 193–213.

[2] Bulmer M.G., The Mathematical Theory of Quantitative Genetics, 2nd ed., Clarendon Press, Oxford, 1985.

[3] Chevalet C., Gillois M., Inbreeding depression and heterosis: expected means and variances among inbred lines and their crosses, Ann. Genet. Sel. Anim. 10 (1978) 73–98.

[4] Cheverud J.M., Routman E.J., Epistasis and its contribution to genetic variance components, Genetics 139 (1995) 1455–1461.

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

[6] Cockerham C.C., Effects of linkage on the covariances between relatives, Genetics 41 (1956) 138–141.

[7] Cockerham C.C., Partition of hereditary variance for various genetic models, Genetics 44 (1959) 1141–1148.

[8] Cockerham C.C., Higher order probability functions of identity of alleles by descent, Genetics 69 (1971) 235–246.

[9] Cockerham C.C., Tachida H., Permanency of response to selection for quantitative characters in finite populations, Proc. Nat. Acad. Sci. USA 85 (1988) 1563–1565.

[10] Comstock R.E., Enfield F.D., Gene number estimation when multiplicative genetic effects are assumed. Growth in flour beetles and mice, Theor. Appl. Genet. 59 (1981) 373–379.

[11] Crow J.F., Basic Concepts in Population, Quantitative and Evolutionary Genetics, W.H. Freeman and Co, New York, 1986.

[12] De Jong G., Van Noordwijk A.J., Acquisition and allocation of resources: genetic (co)variances, selection and life histories, Am. Nat. 139 (1992) 749–770.

[13] De Laguerie P., Olivieri I., Atlan A., Gouyon P.H., Analytic and simulation models predicting positive genetic correlations between traits linked by trade offs, Evol. Ecology 5 (1991) 361–369.

[14] Eshed Y., Zamir D., Less than additive epistatic interactions of quantitative trait loci in tomato, Genetic 143 (1996) 1807–1817.

[15] Fisher R.A., The correlation between relatives on the supposition of Mendelian inheritance, Trans. Roy. Soc. Edinburgh 52 (1918) 399–433.

[16] Gallais A., Covariances between arbitrary relatives with linkage and epistasis in the case of linkage disequilibrium, Biometrics 80 (1974) 429–446.

[17] Gallais A., Théorie de la sélection en amélioration des plantes, Masson, Paris, 1989.

[18] Gillois M., La relation d'identité en génétique, thèse faculté des sciences, Paris, 1964.

[19] Gimelfarb A., Genotypic variation for a quantitative character maintained under stabilizing selection without mutations: epistasis, Genetics 123 (1989) 217–227.

[20] Goodnight C.J., Epistasis and the effect of founder events on the additive genetic variance, Evolution 42 (1988) 441–454.

[21] Grafius J.E., A geometry for plant breeding, Crop Sci. 4 (1964) 241–246.

[22] Harris D.L., Genotypic covariances between inbred relatives, Genetics 50 (1964) 1320–1348.

[23] Hill W.G., Variation in response to selection, in: Pollak E., Kempthorne O., Bailey T.B. (Eds.), Proceedings of the International Conference on Quantitative Genetics, Iowa State University Press, Ames, 1977, pp. 343–345.

[24] Horner T.W., Comstock R.E., Robinson H.F., Non allelic interactions and the interpretation of quantitative genetic data, North Carolina Agricultural Experiment Station Tech. Bull. 118 (1955) 177.

[25] Kempthorne O., The correlation between relatives in a random mating population, Proc. Roy. Soc. London B143 (1954) 103–113.

[26] Kempthorne O., An Introduction to Genetic Statistics, Wiley, New York, 1957.

[27] Kimura M., Maruyama T., The mutational load with epistatic gene interactions in fitness, Genetics 51 (1966) 1337–1351

[28] Langlois B., Incidence de la sélection et de l'homogamie sur les paramètres du modèle additif, Genet. Sel. Evol. 22 (1990) 119–132.

[29] Lin H.X., Qian H.R., Zhuang J.Y., Lu J., Min S.K., Xiong Z.M., Huang N., Zheng K.L., RFLP mapping of QTLs for yield and related characters in rice (Oriza sativa L.), Theor. Appl. Genet. 92 (1996) 920–927.

[30] Malecot G., Les mathématiques de l'hérédité, Masson et Cie, Paris, 1948.

[31] Mansur L.M., Orf J.H., Chase K., Jarvik T., Cregan P.B., Lark K.G., Genetic mapping of agronomic traits using recombinant inbred lines of soybean, Crop Sci. 36 (1996) 1327–1336.

[32] Mather K., Jinks J.L., Biometrical Genetics, Chapman and Hall, London, 1982.

[33] Melchinger A.E., Singh M., Link W., Utz H.F., Von Kittlitz E., Heterosis and gene effects on multiplicative characters. Theoretical relationships and results from Vicia faba L., Theor. App. Genet. 88 (1993) 343–348.

[34] Paterson A.H., Damon S., Hewitt J.D., Zamir D., Rabinowitch H.D., Lincoln S.E., Lander E.S., Tanksley S.D., Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments, Genetics 127 (1991) 191–197.

[35] Powers L., Determining scales and the use of transformations in studies on weight per locule of tomato fruit, Biometrics 6 (1950) 145–163.

[36] Schnell F.W., The covariance betwen relatives in the presence of linkage, in: Hanson W.D. (Ed.), Statistical Genetics and Plant Breeding, Natl. Acad. Sci. Res. Council Pub. 982, 1963, pp. 468–483.

[37] Schnell F.W., Cockerham C.C., Multiplicative vs arbitrary gene action in heterosis, Genetics 131 (1992) 461–463.

[38] Seyffert W., Forkmann G., Simulation of quantitative characters by genes with biochemically definable action VII. Observation and discussion of nonlinear relationships, in: Karlin S., Nevo E. (Eds.), Population Genetics and Ecology, Academic Press, New York, 1976, pp. 431–440.

[39] Tachida H., Cockerham C.C., A building block model for quantitative genetics, Genetics 121 (1989) 839–844.

[40] Turelli M., Barton N.H., Dynamics of polygenic characters under selection, Theor. Pop. Biol. 38 (1990) 1–57.

[41] Van Noordwijk A.J., De Jong G., Acquisition and allocation of resources: their influence on variation in life history tactics, Am. Nat. 128 (1986) 137–142.

[42] Weir B.S., Cockerham C.C., Group inbreeding with two linked loci, Genetics 63 (1969) 711–742.

[43] Weir B.S., Cockerham C.C., Two locus theory in quantitative genetics, in: Pollak E., Kempthorne O., Bailey T.B. (Eds.), Proceedings of the International Conference on Quantitative genetics, Iowa State University Press, Ames, 1977, pp. 247–269.

[44] Whitlock M.C., Phillips P.C., Wade M.J., Gene interaction affects the additive genetic variance in subdivided populations with migration and extinction, Evolution 47 (1993) 1758–1769.

[45] Wright S., The analysis of variance and the correlations between relatives with respect to deviation from an optimum, J. Genet. 30 (1935) 243–256.

[46] Zhao-Bang Z., Cockerham C.C., Variance of neutral genetic variances within and between populations for a quantitative character, Genetics 129 (1991) 535–553.

APPENDIX A: Decomposition of the genotypic value under MGA according to the factorial method

Additive effects

From equation (3), and assuming linkage equilibrium

$$\mu_{ij_s} = a_{ij_s} \prod_{t \neq s} \bar{a}_t \tag{A.1}$$

In the factorial method, the additive effect of allele i at locus s is defined as

$$\alpha_{i_s} = \left(\sum_j p_{j_s} \mu_{ij_s} \right) - \mu \tag{A.2}$$

where symbols are the same as in the text. Using the expression (3) for μ and factoring $\prod_{t \neq s} \bar{a}_t$, one obtains the expression (6b) for the additive effect of allele i at locus s under MGA.

Dominance effects

The dominance effect β_{ij_s} between alleles i and j at locus s is defined classically as

$$\beta_{ij_s} = \mu_{ij_s} - \alpha_{i_s} - \alpha_{j_s} - \mu \tag{A.3}$$

The expression (7b) is obtained by using formulae (A.1) for μ_{ij_s} and (6b) for α_{i_s} and α_{j_s}, and again factoring $\prod_{t \neq s} \bar{a}_t$.

Epistatic effects

As pointed out in the text, the factorial decomposition applied under MGA generates epistatic effects. The additive by additive effect $(\alpha\alpha)_{i_s k_t}$ pertaining to allele i at locus s and allele k at locus t is defined as:

$$(\alpha\alpha)_{i_s k_t} = \mu_{i+_s k+_t} - \alpha_{i_s} - \alpha_{k_t} - \mu \tag{A.4}$$

where $\mu_{i+_s k+_t}$ is the expectation of genotypic values for individuals having received gene i at locus s and gene k at locus t from one of their parents (e.g. sire), the genes transmitted by the other parent being any gene drawn at random in the population. Under linkage equilibrium

$$\mu_{i+_s k+_t} = \sum_j \sum_l p_{j_s} p_{l_t} a_{ij_s} a_{kl_t} \left(\prod_{u \neq s,t} \bar{a}_u \right) \tag{A.5}$$

Using the expression for $\alpha^*_{i_s}$ given in equation (6a) and putting it into equations (A.4 and A.5) gives

$$(\alpha\alpha)_{i_s k_t} = \left[\left(\alpha^*_{i_s} + \bar{a}_s \right) \left(\alpha^*_{k_t} + \bar{a}_t \right) - \alpha^*_{i_s} \bar{a}_t - \alpha^*_{k_t} \bar{a}_s - \bar{a}_t \bar{a}_s \right] \left(\prod_{u \neq s,t} \bar{a}_u \right) \tag{A.6}$$

which reduces to equation (8) after rearranging.

APPENDIX B: Partition of the genotypic variance under MGA

Consider the same decomposition as in equation (1) with realized values replaced by random variables pertaining to the same genetic effects defined for a random by chosen individual in the population

$$g_z = \mu + \sum_s (\alpha_{1_s} + \alpha_{2_s})$$

$$+ \sum_s \beta_s$$

$$+ \sum_s \sum_{t>s} \left[(\alpha\alpha)_{1_s1_t} + (\alpha\alpha)_{1_s2_t} + (\alpha\alpha)_{2_s1_t} + (\alpha\alpha)_{2_s2_t} \right] \quad \text{(B.1)}$$

$$+ \ldots$$

where the symbols i and j coding for the alleles are replaced by the integers 1 and 2 designating the genes transmitted by the sire and dam, respectively (those figures being omitted in dominance effects for the sake of simplicity). The same assumptions are made as before (i.e. infinite population size, linkage equilibrium and panmixia), resulting in orthogonal decomposition with independent random variables.

Additive genetic variance (σ_A^2)
By definition

$$\sigma_A^2 = \sum_s \text{Var}(\alpha_{1_s}) + \text{Var}(\alpha_{2_s}) \quad \text{(B.2)}$$

and, because the paternal and maternal components are playing the same roles, $\text{Var}(\alpha_{1_s}) = \text{Var}(\alpha_{2_s}) = \sum_i p_{i_s}\alpha_{i_s}^2$ and $\sigma_A^2 = 2\sum_{s=1}^{S}\left(\sum_i p_{i_s}\alpha_{i_s}^2\right)$. The additive genetic variance (10a) is then obtained by using the expression of α_{i_s} in equation (6b), and by setting

$$\sigma_{a_s}^2 = 2\sum_i p_{i_s}\alpha_{i_s}^{*\,2} \quad \text{(B.3)}$$

for the additive variance among $\alpha_{i_s}^*$ values at locus s.

Dominance genetic variance (σ_D^2)

Similarly, $\sigma_D^2 = \sum_s \text{Var}(\beta_s)$. Knowing that $\beta_{ij_s} = \beta_{ij_s}^*\left(\prod_{s\neq t}\bar{a}_t\right)$, and letting, as before

$$\sigma_{d_s}^2 = \sum_i\sum_j p_{i_s}p_{j_s}\beta_{ij_s}^{*\,2} \quad \text{(B.4)}$$

one obtains the expression (11a) for the dominance genetic variance.

Additive by additive genetic variance
From (B.1)

$$\sigma_{AA}^2 = \sum_s\sum_{t>s}\left[\text{Var}\big((\alpha\alpha)_{1_s1_t}\big) + \text{Var}\big((\alpha\alpha)_{1_s2_t}\big)\right.$$

$$\left. + \text{Var}\big((\alpha\alpha)_{2_s1_t}\big) + \text{Var}\big((\alpha\alpha)_{2_s2_t}\big)\right] \quad \text{(B.5)}$$

As previously, paternal and maternal contributions are equivalent, and the four elementary variances are equal. Thus (omitting subscripts for parental contributions) $\sigma_{AA}^2 = 4 \sum_s \sum_{t>s} \text{Var}\left((\alpha\alpha)_{st}\right)$, with $\text{Var}\left((\alpha\alpha)_{st}\right) = \sum_i \sum_k p_{i_s} p_{k_t} (\alpha\alpha)_{i_s k_t}^2$. Now, using equation (8) for $(\alpha\alpha)_{i_s k_t}$ and equation (B.3) for the relationship between $\alpha_{i_s}^*$ and $\sigma_{a_s}^2$, we have

$$\text{Var}\left((\alpha\alpha)_{st}\right) = \frac{1}{4} \sigma_{a_s}^2 \sigma_{a_t}^2 \left(\prod_{u \neq s,t} \bar{a}_u^2 \right)$$

Finally,

$$\sigma_{AA}^2 = \sum_{s=1}^{S} \sum_{t>s} \left[\sigma_{a_s}^2 \sigma_{a_t}^2 \left(\prod_{u \neq s,t} \bar{a}_u^2 \right) \right] \tag{B.6}$$

Other variance components

Additive by dominance genetic variance, dominance by dominance genetic variance, as well as variances pertaining to higher order interactions can be derived in the same way. For instance, the expression for additive by additive epistatic variance (σ_{AAA}^2) can be obtained along the same lines as in equation (B.6). In that case, there are eight variance terms for $(\alpha\alpha\alpha)_{stu}$ paternal and maternal contributions which are equal. As $\sum_i p_{i_s} (\alpha_{i_s}^*)^2 = \sigma_{a_s}^2/2$, this factor 8 cancels out with $(1/2)^3$ owing to the introduction of the product $\sigma_{a_s}^2 \sigma_{a_t}^2 \sigma_{a_u}^2$. As there are 3! possible permutations of s, t and u in $(\alpha\alpha\alpha)_{stu}$, which are equivalent, the final expression for σ_{AAA}^2 is obtained by summing up elementary contributions over different s, t and u loci and by dividing by 3! (see equation 12d).