Théorie approchée de la sélection pour un caractère dû à un nombre fini de locus

Claude C. Chevalet

Original article

An approximate theory of selection assuming a finite number of quantitative trait loci

C Chevalet

Institut National de la Recherche Agronomique,
Centre de Recherches de Toulouse,
Laboratoire de Génétique Cellulaire,
BP 27, 31326 Castanet-Tolosan cedex, France

Summary – An approximate theory of mid-term selection for a quantitative trait is developed for the case when a finite number of unlinked loci contribute to phenotypes. Assuming Gaussian distributions of phenotypic and genetic effects, the analysis shows that the dynamics of the response to selection is defined by one single additional parameter, the effective number L_e of quantitative trait loci (QTL). This number is expected to be rather small (3–20) if QTLs have variable contributions to the genetic variance. As is confirmed by simulation, the change with time of the genetic variance and of the cumulative response to selection depend on this effective number of QTLs rather than on the total number of contributing loci. The model extends the analysis of Bulmer, and shows that an equilibrium structure arises after a few generations in which some amount of genetic variability is hidden by gametic disequilibria. The additive genetic variance V_A and the genic variance V_a remain linked by: $V_A \simeq V_a - \kappa(1 - 1/L_e)h^2 V_A$, where κ is the proportion of variance removed by selection, and h^2 the current heritability of the trait. From this property, a complete approximate theory of selection can be developed, and modifications of correlations between relatives can be proposed. However, the model generally overestimates the cumulative response to selection except in early generations, which defines the time scale for which the present theory is of potential practical value.

quantitative genetics / selection / genetic variance

Résumé – **Théorie approchée de la sélection pour un caractère dû à un nombre fini de locus.** *Une théorie approchée de la sélection est développée dans le cas d'un caractère quantitatif dont la variabilité génétique est due à un nombre fini de locus génétiquement indépendants. Le calcul est développé analytiquement en admettant que toutes les distributions statistiques peuvent être approchées par des lois normales. L'analyse montre que le comportement global du système génétique dépend essentiellement d'un «nombre efficace de locus», L_e, dont les valeurs vraisemblables sont sans doute faibles*

(3 à 20). Des simulations confirment le rôle de ce paramètre pour caractériser la réponse cumulée à la sélection et la structure génétique de la population. Le modèle généralise l'analyse de M Bulmer. Après quelques générations d'un régime de sélection, une fraction de la variance génétique reste « cachée» sous la forme de covariances négatives, de sorte que la variance génétique additive V_A et la variance génique V_a demeurent liées par la relation : $V_A \simeq V_a - \kappa(1 - 1/L_e)h^2 V_A$, où κ est la fraction de variance réduite par la sélection, et h^2 est l'héritabilité actuelle du caractère. Cette structuration de la variance génétique sous sélection permet de proposer des expressions modifiées des covariances entre apparentés issus de parents sélectionnés, et de développer une théorie complète de la sélection. Sauf à court et moyen terme, les prédictions quantitatives sont surestimées par le modèle gaussien, ce qui délimite le champ d'application pratique de la théorie.

génétique quantitative / sélection / variance génétique

INTRODUCTION

Models of quantitative genetics are generally developed under the assumptions of the infinitesimal model, which states that a very large number of genetically unlinked loci contribute to the genetic variance of a trait. More precisely, it is assumed that all contributions of individual loci are of the same order of magnitude. This hypothesis ensures that the distribution of breeding values is Gaussian, and validates the whole statistical apparatus that made the statistical developments of applied quantitative genetics possible and its practical achievements. The scope of the present paper is to develop an approximate theory that may cope with more general genetic situations, owing to the introduction of an additional parameter characterizing a quantitative trait. The cases considered in the following involve variable contributions to the quantitative trait of a finite number of genetically unlinked loci. The derivations rely on the hypothesis that all distributions can be approximated by a Gaussian, following the method illustrated by Lande (1976) and Chevalet (1988). This makes it possible to define an analytical theory of selection with a model that seems less unrealistic than the usual infinitesimal hypothesis involving very many unlinked quantitative loci.

Two main qualitative predictions are derived from the Gaussian model: (i) a single parameter, which can be called the effective number of quantitative trait loci (QTL), is a good summary of the distribution of the variable contributions of QTLs to the genetic variance of the trait; and (ii) under continued selection the amount of genetic variability that is hidden by negative correlations between contributions of different loci can be calculated as a function of selection intensity, the current additive genetic variance, and the effective number of QTLs. In addition to analytical derivations, simulations were performed in order to evaluate the qualitative and quantitative importance of departures from normality.

GENETIC MODEL

We consider a diploid monoecious population of N reproducing individuals per generation, with L loci. Let

$$\mathbf{g}^{(m)'} = (g_1^{(m)}, g_2^{(m)}, \cdots, g_L^{(m)})$$
$$\mathbf{g}^{(f)'} = (g_1^{(f)}, g_2^{(f)}, \cdots, g_L^{(f)})$$

be the genotypes of a male and a female gamete, respectively. The numbers $g_\ell^{(m)}$ and $g_\ell^{(f)}$ are defined as absolute effects of the genes carried by the corresponding loci ℓ. These effects are distributed in the population, and their joint distribution is assumed to be multivariate normal. Assuming symmetry between male and female contributions, any value is written in the following way:

$$\mathbf{g}^{(m)} = \overline{\mathbf{g}} + \boldsymbol{\gamma}^{(m)}$$
$$\mathbf{g}^{(f)} = \overline{\mathbf{g}} + \boldsymbol{\gamma}^{(f)}$$

where $\overline{\mathbf{g}}$ is the mean value of a gamete and $\boldsymbol{\gamma}$ a residual.

Matings are assumed to be random, so that the variance covariance matrix of gene effects in new zygotes takes the form

$$\mathcal{G} = \mathrm{Cov}\left[\begin{pmatrix}\mathbf{g}^{(m)} \\ \mathbf{g}^{(f)}\end{pmatrix}, (\mathbf{g}^{(m)'}, \mathbf{g}^{(f)'})\right]$$

$$\mathcal{G} = \begin{pmatrix} \mathbf{G} & 0 \\ 0 & \mathbf{G} \end{pmatrix} \tag{1}$$

where $\mathbf{G} = \mathrm{Cov}(\mathbf{g}^{(m)}, \mathbf{g}^{(m)'}) = \mathrm{Cov}(\mathbf{g}^{(f)}, \mathbf{g}^{(f)'})$ is the variance covariance matrix between gene effects of a gamete drawn from the reproducing individuals in the preceding generation.

The value of phenotype P in a zygote with $(\mathbf{g}^{(m)}, \mathbf{g}^{(f)})$ genotypic value is assumed to depend linearly on gene effects:

$$P(\mathbf{g}^{(m)}, \mathbf{g}^{(f)}) = \mathbf{B}' \cdot \overline{\mathbf{g}} + \mathbf{B}' \cdot \overline{\mathbf{g}} + \mathbf{B}' \cdot \boldsymbol{\gamma}^{(m)} + \mathbf{B}' \cdot \boldsymbol{\gamma}^{(f)} + \mathbf{E} \tag{2}$$

where $\mathbf{B}$ is a $(L \times 1)$ vector. Note that considering several vectors $\mathbf{B}$ allows several traits to be considered simultaneously.

The genotypic distribution among the zygotes is given by equation [1], so that the first 2 moments of a trait P are:

$$\overline{P} = 2\mathbf{B}' \cdot \overline{\mathbf{g}} \tag{3}$$
$$\mathrm{Var}(P) = 2\mathbf{B}'\mathbf{G}\mathbf{B} + V_E \tag{4}$$

where $(2\mathbf{B}'\mathbf{G}\mathbf{B})$ is the additive genetic variance V_A of the trait, and V_E is the variance of environmental effects on the trait. Similarly, the genetic covariance between P and a second trait Q characterized by a vector $\mathbf{C}$, is: $\mathrm{Cov}(P, Q) = 2\mathbf{C}'\mathbf{G}\mathbf{B}$.

Under the Gaussian approximation, the genetic modifications induced by selection on the phenotype are calculated from the regression equations, and depend only on the first 2 moments of the phenotypic changes. Thus, the exact selection rule is not important. For example, truncation selection and stabilizing selection with a Gaussian fitness function yield the same predictions provided they are characterized by the same changes in the mean and variance of phenotypes. The relevant parameters are defined as follows, where subscript s refers to values after selection:

– the selection intensity $\imath$, relating the change in mean phenotypic value to the phenotypic standard deviation

$$\overline{P}_s - \overline{P} = \imath\sqrt{\mathrm{Var}(P)}$$

– the relative change in the variance κ

$$\mathrm{Var}_s(P) = (1 - \kappa)\mathrm{Var}(P)$$

Assuming that selection occurs among a large population of zygotes, the values of covariances between gene effects in the selected individuals is:

$$\mathrm{Cov}_s(\boldsymbol{\gamma}_\ell^{(m)}, \boldsymbol{\gamma}_j^{(m)}) = \mathrm{Cov}_s(\boldsymbol{\gamma}_\ell^{(f)}, \boldsymbol{\gamma}_j^{(f)}) = \mathbf{G}_{\ell j} - \kappa\frac{K_\ell K_j}{\mathrm{Var}(P)} \qquad [5]$$

$$\mathrm{Cov}_s(\boldsymbol{\gamma}_\ell^{(m)}, \boldsymbol{\gamma}_j^{(f)}) = \mathrm{Cov}_s(\boldsymbol{\gamma}_\ell^{(f)}, \boldsymbol{\gamma}_j^{(m)}) = -\kappa\frac{K_\ell K_j}{\mathrm{Var}(P)} \qquad [6]$$

where K_ℓ is defined as

$$K_\ell = \sum_{j=1}^{j=L} G_{\ell j} B_j \qquad [7]$$

Then, taking account of gametogenesis, and $r_{\ell j}$ being the recombination fraction between loci ℓ and j, recurrence relationships between 2 successive generations (t) and $(t+1)$ can be derived for the mean and the variance covariance matrix of gene effects (Lande, 1976; Chevalet, 1988):

– mean effects $\overline{g}'$s:

$$E(\overline{g}_\ell^{(t+1)}) = E(\overline{g}_\ell^{(t)}) + \frac{\imath}{\sqrt{\mathrm{Var}_{(t)}(P)}}K_\ell^{(t)} \qquad [8]$$

– within population structure:

$$G_{\ell j}^{(t+1)} = (1 - \frac{1}{2N} - r_{\ell j})G_{\ell j}^{(t)} - (1 - \frac{1}{N})\frac{\kappa K_\ell^{(t)} K_j^{(t)}}{\mathrm{Var}_{(t)}(P)} \qquad [9]$$

– variance of the mean values (drift effect):

$$\mathrm{Cov}(\overline{g}_\ell^{(t+1)}, \overline{g}_j^{(t+1)}) = \mathrm{Cov}(\overline{g}_\ell^{(t)}, \overline{g}_j^{(t)}) + \frac{1}{N}\left(\frac{1}{2}G_{\ell j}^{(t)} - \frac{\kappa K_\ell^{(t)} K_j^{(t)}}{\mathrm{Var}_{(t)}(P)}\right) \qquad [10]$$

SIMULATION MODEL

The simulated model shares the same general hypotheses as the analytical scheme (same initial value of heritability, same distribution of the contributions of loci to the genetic variance), but is a completely discrete genetic model.

At each locus, a finite number of alleles are assigned additive effects that sum up to the breeding value of a zygote, to which a Gaussian random variable is added to simulate the environmental effect. The additive effects of alleles are drawn in the initial generation from a Gaussian distribution, and adjusted to yield the specified heritability and distribution of contributions among loci. The population size is described by 2 numbers: the number of zygotes; and the number N of selected adults. Truncation selection on individual phenotypic values is performed, and adults are mated at random (with selfing occurring with a probability of $1/N$). The genetic make-up of gametes produced by the parents are generated using a pseudo-random-number generator to simulate Mendelian segregations.

Programs allow for various initial distributions of allelic effects within and across loci, several selection rules (truncation selection is used here), and various linkage relationships between loci (fixed at $1/2$ in the present work). Outputs from the program include, at each generation, the mean values and standard deviations over replicated runs of the following criteria: mean breeding value; genetic and genic variances, effective numbers of loci (equations [17] and [18] below) and of alleles per locus; mean homozygosity; proportion of fixed loci; and (for models assuming independent loci) the T parameter defined in the following (equation [21]). One-hundred runs were performed for each considered case. Programs were written in Fortran 77 and were run on a UNIX machine.

ANALYTICAL DERIVATIONS

The effective number of QTLs

With equal contributions of unlinked loci, equation [9] leads to only 2 equations describing the change with time of 2 macroscopic statistics, the additive genetic variance V_A , and the genic variance V_a (ie the sum of the variances contributed by the loci). Removing time indices (the asterisk denoting the next generation), the equations are (Chevalet, 1988):

$$V_A^* = \frac{1}{2}(1 - \frac{1}{N})(1 - \kappa h^2)V_A + \frac{1}{2}V_a \qquad [11]$$

$$V_a^* = (1 - \frac{1}{2N})V_a - (1 - \frac{1}{N})\kappa h^2 \frac{V_A}{2L} \qquad [12]$$

where h^2 is the current value of heritability, $h^2 = \dfrac{V_A}{\text{Var}(P)}$. The genic variance V_a can be written as:

$$V_a = 2\sum_{\ell=1}^{\ell=L} G_{\ell\ell}B_\ell^2 = V_A - D$$

D being the sum of the contributions to V_A of the covariances between gene effects at different loci.

In the case of unlinked loci, equation [9] has 2 types, for diagonal terms ($r_{\ell j} = 0$) and for non-diagonal terms ($r_{\ell j} = \frac{1}{2}$). Multiplying equation [9] by B_ℓ and summing over ℓ yields:

$$\sum_{\ell=1}^{\ell=L} B_\ell G_{\ell j}^* = (1 - \frac{1}{2N}) \sum_{\ell=1}^{\ell=L} B_\ell G_{\ell j} - \frac{1}{2} \left(\sum_{\ell=1}^{\ell=L} B_\ell G_{\ell j} - B_j G_{jj} \right)$$

$$- (1 - \frac{1}{N}) \frac{\kappa K_j}{\mathrm{Var}(P)} \sum_{\ell=1}^{\ell=L} B_\ell K_\ell$$

$$K_j^* = \frac{1}{2}(1 - \frac{1}{N}) K_j + \frac{1}{2} B_j K_j - (1 - \frac{1}{N}) \frac{\kappa K_j}{\mathrm{Var}(P)} \frac{V_A}{2}$$

thus:

$$K_j^* = \frac{1}{2}(1 - \frac{1}{N}) \left(1 - \frac{\kappa V_A}{\mathrm{Var}(P)} \right) K_j + \frac{1}{2} B_j G_{jj} \qquad [13]$$

Multiplying equation [13] by B_j and summing yields equation [11], as in the case of uniform contributions. In contrast, summing the diagonal products $B_j^2 G_{jj}^*$ in equation [9] gives:

$$\sum_{j=1}^{j=L} B_j^2 G_{jj}^* = (1 - \frac{1}{2N}) \sum_{j=1}^{j=L} B_j^2 G_{jj} - (1 - \frac{1}{N}) \frac{\kappa}{\mathrm{Var}(P)} \sum_{j=1}^{j=L} B_j^2 K_j^2$$

$$\frac{1}{2} V_a^* = (1 - \frac{1}{2N}) \frac{1}{2} V_a - (1 - \frac{1}{N}) \frac{\kappa}{\mathrm{Var}(P)} \sum_{j=1}^{j=L} B_j^2 K_j^2 \qquad [14]$$

Introducing deviations X_j (resp Y_j) of the contributions $B_j K_j$ (resp $G_{jj} B_j^2$) of locus j from the mean contribution $\frac{1}{2L} V_A$ (resp $\frac{1}{2L} V_a$) of a locus

$$X_j = B_j K_j - \frac{1}{L} \frac{V_A}{2}$$

$$Y_j = G_{jj} B_j^2 - \frac{1}{L} \frac{V_a}{2}$$

gives:

$$\sum_{j=1}^{j=L} B_j^2 K_j^2 = \sum_{j=1}^{j=L} \left(X_j + \frac{V_A}{2L} \right)^2 = \frac{V_A^2}{4L} + \sum_{j=1}^{j=L} X_j^2$$

Equation [14] becomes

$$V_a^* = (1 - \frac{1}{2N})V_a - (1 - \frac{1}{N})\frac{\kappa}{2L\mathrm{Var}(P)}\left(V_A^2 + 4L\sum_{j=1}^{j=L}X_j^2\right)$$

which can be written in a form similar to equation [12]:

$$V_a^* = (1 - \frac{1}{2N})V_a - (1 - \frac{1}{N})\frac{\kappa}{2}\frac{V_A^2}{L_e}\frac{1}{\mathrm{Var}(P)} \tag{15}$$

defining the effective number L_e of quantitative trait loci as:

$$L_e = \frac{LV_A^2}{V_A^2 + 4L\sum_{j=1}^{j=L}X_j^2} \tag{16}$$

this can also be written in the following forms:

$$L_e = \frac{\left(\sum_{j=1}^{j=L}B_jK_j\right)^2}{\left(\sum_{j=1}^{j=L}B_j^2K_j^2\right)} \tag{17}$$

$$\frac{L}{L_e} = 1 + CV^2 \tag{18}$$

where CV is defined as the coefficient of variation of the contributions of the various loci to the total additive genetic variance

$$CV^2 = \frac{\frac{1}{L}\sum_{j=1}^{j=L}X_j^2}{\left(\frac{1}{2}V_A\right)^2}$$

In addition to the 2 main equations [11] and [14], the following equations for the deviations X_j and Y_j can be derived:

$$X_j^* = \frac{1}{2}(1 - \frac{1}{N})(1 - \frac{\kappa V_A}{\mathrm{Var}(P)})X_j + \frac{1}{2}Y_j$$

$$Y_j^* = (1 - \frac{1}{2N})Y_j - (1 - \frac{1}{N})\frac{\kappa}{\mathrm{Var}(P)}\left(X_j^2 + \frac{V_A X_j}{L} - \frac{1}{L}\sum_{\ell=1}^{\ell=L}X_\ell^2\right)$$

It can seen that these deviations would remain null if they are so at some time. However, it would be interesting to check if this null state is stable with respect

to perturbations. Together with equations [12], [15] and [16], these equations form a closed set of $2L$ independent equations which can be extracted from the set of $L(L+1)/2$ (equation [9]). This exact result, which exhibits a hierarchical structure within the system [9], is completed by the approximate result that only 2 equations are needed to get a comprehensive description of the dynamics of the system. In fact, the value of L_e, as defined above, depends on time unless initial conditions are such that $L_e = L$. Various numerical calculations comparing the change with time of V_A, either from full equations [9] or from simple equations [11] and [12], with the proper initial value of L_e, show that for many generations no significant discrepancy can be found. As far as only macroscopic parameters are of interest (genetic variance or response to selection on the phenotypic scale), it seems valuable to simplify the complete system, and reduce its description to both equations [11] and [14], where L_e is related to the microscopic (unobservable) parameters by equations [16]–[18].

EQUILIBRIUM STRUCTURE UNDER SELECTION (BULMER EFFECT)

Directional selection for a trait due to the additive effects of several loci develops negative correlations between the contributions of distinct loci. In the statistical setting of the infinitesimal model, in which loci are not individually considered, this effect has been proven by Bulmer (1971) by considering the regression of the genotypic value on phenotypes after selection. In a very large population, and assuming initial linkage equilibrium, he derived the following recursion (a special case of equations [11] and [12]):

$$V_A^{(t+1)} = \frac{1}{2}(1 - \kappa \frac{V_A^{(t)}}{\text{Var}_{(t)}(P)})V_A^{(t)} + \frac{1}{2}V_a^{(t)}$$
$$V_a^{(t+1)} = V_a^{(t)}$$

He also showed that after a few generations, an equilibrium structure arises, in which the genic variance V_a remains equal to the initial genetic variance $V_A^{(0)}$ and the genetic variance is fixed at a reduced value dependent on selection strength. The limit values are such that

$$V_A^{(\infty)} - V_a^{(\infty)} = -\kappa \frac{\left(V_A^{(\infty)}\right)^2}{\text{Var}_\infty(P)} \qquad [19]$$
$$V_a^{(\infty)} = V_A^{(0)}$$

Equation [19] gives the total amount contributed at equilibrium by negative correlations (*ie* linkage disequilibria) to the genetic variance.

In the first generation, this result can be shown directly by a genetic analysis, under the hypothesis of the infinitesimal model, starting from a model involving multiallelic distributions if the initial population is assumed to be in Hardy–Weinberg equilibrium at all loci, and in linkage equilibrium for all pairs of loci. A more general treatment of the problem is proposed by Turelli and Barton (1990),

based on the calculations of all the moments of distributions. However, unless special hypotheses are stated, their approach does not provide explicit recurrence relationships after the first generation.

In the present model, the genetic variance decreases to zero as soon as L is finite when selection is active (κ is positive), and if N is finite selection accelerates the fixation process (Chevalet, 1988). However, a qualitative property similar to Bulmer's result still holds: under continuous selection (constant selection strength), the following approximate relationship holds at any generation t after 4 or 5 generations under the same selection rules:

$$V_A^{(t)} - V_a^{(t)} \simeq -\kappa(1 - \frac{1}{L_e})\frac{\left(V_A^{(t)}\right)^2}{\text{Var}_{(t)}(P)} \qquad [20]$$

This shows that, while genetic variances decrease to zero, the total contribution of negative correlations remains proportional to the square of the available genetic variance.

The result is obtained by introducing (for $\kappa \neq 0$) a new variable $T^{(t)}$:

$$V_a^{(t)} = V_A^{(t)} + \kappa\frac{\left(V_A^{(t)}\right)^2}{\text{Var}_{(t)}(P)}T^{(t)} \qquad [21]$$

and rewriting equations [11] and [15], with the 2 variables V_a and T. Writing equation [21] as:

$$V_A^{(t)} = F(V_a^{(t)}, T^{(t)}) \qquad [22]$$

the recursion in T is obtained as follows (discarding time indices as before):

$$T^* = \frac{V_a^* - V_A^*}{\kappa\left(V_A^*\right)^2/\text{Var}^*(P)}$$

The numerator can be written as:

$$V_a^* - V_A^* = \frac{1}{2}(1 - \frac{1}{N})\left(V_a - V_A + (1 - \frac{1}{L_e})\kappa\frac{(V_A)^2}{\text{Var}(P)}\right)$$

$$= \frac{1}{2}(1 - \frac{1}{N})\kappa\frac{(V_A)^2}{\text{Var}(P)}(T + 1 - \frac{1}{L_e})$$

In the denominator, V_A^* is written as

$$V_A^* = \frac{1}{2}V_a + \frac{1}{2}(1 - \frac{1}{N})V_A(1 - \kappa\frac{V_A}{\text{Var}(P)})$$

$$= \frac{1}{2}(V_A + \kappa\frac{(V_A)^2}{\text{Var}(P)}T) + \frac{1}{2}(1 - \frac{1}{N})V_A(1 - \kappa\frac{V_A}{\text{Var}(P)})$$

$$V_A^* = V_A\left(1 - \frac{1}{2N} + \frac{1}{2}\kappa\frac{V_A}{\text{Var}(P)}\left(T - (1 - \frac{1}{N})\right)\right) \qquad [23]$$

thus:

$$T^* = \frac{1}{2}(1 - \frac{1}{N})\frac{T + 1 - \frac{1}{L_e}}{\left(1 - \frac{1}{2N} + \frac{1}{2}\kappa\frac{V_A}{\text{Var}(P)}\left(T - (1 - \frac{1}{N})\right)\right)^2}\frac{\text{Var}^*(P)}{\text{Var}(P)} \qquad [24]$$

The recursion in (V_a, T) can then be derived using function F (equation [22]) and assuming either that the phenotypic variance is constant ($\text{Var}(P) = V_P$), or that the environmental variance (V_E) is constant. In the latter case $\text{Var}(P)$ and $\text{Var}^*(P)$ are written as $F(V_a, T) + V_E$ and $V_A^* + V_E$ using expressions [22] and [23].

In the case of constant phenotypic variance V_P, we obtain the system:

$$V_a^* = (1 - \frac{1}{2N})V_a - (1 - \frac{1}{N})\frac{\kappa}{2L_eV_P}\left(F(V_a, T)\right)^2 \qquad [25]$$

$$T^* = \frac{1}{2}(1 - \frac{1}{N})\frac{T + 1 - \frac{1}{L_e}}{\left(1 - \frac{1}{2N} + \frac{1}{2}\kappa\frac{F(V_a, T)}{V_P}\left(T - (1 - \frac{1}{N})\right)\right)^2} \qquad [26]$$

Written in this way, it can be seen that V_a is a slowly varying expression, for N and L_e not too small, while T reaches the neighborhood of a limit $\overline{T}$ in a few generations:

$$\overline{T} = \frac{(1 - \frac{1}{N})(1 - \frac{1}{L_e})}{1 - \frac{1}{N} + \frac{1}{2N^2}} \qquad [27]$$

This yields equation [20] above. In fact, as is done in *Appendix*, we can show analytically that T reaches the neighborhood of 1 within 4 to 5 generations; after this first step, the convergence to $\overline{T}$ may be rather slow and depends on the relative values of κ, L_e and N (numerical calculations). The same occurs for both models of phenotypic variance (constant phenotypic or environmental variances), with the same limit $\overline{T}$ and the same kind of convergence.

An approximate complete solution

The analysis of the model can be further developed, owing to the reduction to 2 equations, and even to a single equation. Indeed, since T reaches its limit in a few generations, replacing $T^{(t)}$ by $\overline{T}$ in equation [21] or [22] allows $V_A^{(t)}$ to be written as an algebraic function of $V_a^{(t)}$.

Equation [15] becomes:

$$V_a^{(t+1)} = (1 - \frac{1}{2N})V_a^{(t)} - (1 - \frac{1}{N})\frac{\kappa}{2}\frac{[F(V_a^{(t)}, \overline{T})]^2}{L_e \, \text{Var}_{(t)}(P)} \qquad [28]$$

If N and L_e are not too small, and their ratio is finite, this difference equation can be transformed into a differential equation. Assuming a constant phenotypic variance $(\mathrm{Var}_{(t)}(P) = V_P)$, we obtain a scaled time

$$u = \frac{t}{2N} \tag{29}$$

and the following notations:

$$a = \frac{1}{2}\frac{N}{L_e}(1 - \frac{1}{N})\frac{1}{\overline{T}} = \frac{1}{2}\frac{N - 1 + \frac{1}{2N}}{L_e - 1} \tag{30}$$

$$b = \frac{V_P}{\kappa \overline{T}} \tag{31}$$

and then, substituting F by its definition, equation [28] is approximated by

$$\frac{dV_a}{du} = -V_a - ab\left(1 + 2\frac{V_a}{b} - \sqrt{1 + 4\frac{V_a}{b}}\right) \tag{32}$$

Integration gives the (scaled) time $u_2 - u_1$ corresponding to a reduction in the genic variance, from $V_a^{(t_1)}$ to $V_a^{(t_2)}$. The result is easily obtained by changing the variable V_a to $W = F(V_a, \overline{T})$ (W is used here instead of V_A to avoid confusion between the true value of genetic variance and its approximation). The differential equation becomes

$$\frac{dW}{du} = -\frac{W(1 + \frac{1 + 2a}{b}W)}{1 + \frac{2}{b}W} \tag{33}$$

and the solution is

$$u_1 - u_2 = \ln\frac{W_2}{W_1} + \frac{1 - 2a}{1 + 2a}\ln\frac{1 + \frac{1 + 2a}{b}W_2}{1 + \frac{1 + 2a}{b}W_1} \tag{34}$$

which gives W_2 as an implicit function of W_1 and of $(u_2 - u_1)$. Using this approximation, an equation for the cumulated response to selection can be derived. The cumulative response from time t_1 to time t_2 is

$$\overline{P}_2 - \overline{P}_1 = \frac{\imath}{\sqrt{V_P}}\left(V_A^{(t_1)} + V_A^{(t_1+1)} + \ldots + V_A^{(t_2-1)}\right)$$

Changing V_A to its approximation W, and replacing the sum by an integral, we can write:

$$\overline{P}_2 - \overline{P}_1 \simeq \frac{2N\imath}{\sqrt{V_P}}\ I(\frac{t_1}{2N}, \frac{t_2}{2N}) \tag{35}$$

where

$$I(u_1, u_2) = \int_{u_1}^{u_2} W(u)du$$

This integral can be readily calculated by rewriting equation [33] as

$$\frac{dW}{W} + \frac{2}{b}dW + du + \frac{1+2a}{b}W du = 0$$

so that, taking account of equation [34], $I(u_1, u_2)$ is

$$I(u_1, u_2) = \frac{2}{1+2a}(W_1 - W_2) - b\frac{1-2a}{(1+2a)^2} \ln \frac{1 + \dfrac{1+2a}{b}W_1}{1 + \dfrac{1+2a}{b}W_2} \qquad [36]$$

The set of equations [34]–[36] provides an explicit solution to the problem, which can be completed by a further equation giving, from equations [10], the variance of the response due to sampling. Comparing the numerical results obtained with this continuous approximation with that derived from the iterative use of equations [11] and [15] shows a very good agreement, as far as the comparison does not involve the first generations if initial conditions assume linkage equilibrium ($V_A = V_a$). In the latter case, the continuous approximation underestimates the initial response to selection. Although it is derived under the hypothesis that N and L_e are rather large, the approximation is still correct for values of L_e as small as 5.

A similar analysis can be carried out for the model assuming a constant environmental variance, rather than a constant phenotypic variance.

DISCUSSION

The preceding calculations show that the dynamics of the multilocus system considered can be described by introducing a single additional parameter (the effective number of QTLs), as compared to the standard statistical setting of quantitative genetics. The result holds as far as only macroscopic properties of the system are considered, and for a limited number of generations, because the non-linear features of the system (equation [9]) do not allow the derivation of a uniform upper boundary for the deviations of individual locus contributions from their mean values. This parameter, L_e, also allows the structure of genetic variance to be predicted, according to the generalization of Bulmer's result to a finite population and to a finite number of QTLs (equation [20]).

The number of QTLs

Recent results of QTL detection, mainly in plants, suggest a rather small number of loci contributing a significant part to the genetic variance of quantitative traits. This does not mean that a few genes are involved in the make-up of the trait, but that only loci contributing a rather large genetic variance can be detected by segregation analysis. A simple way to describe the distribution of individual

contributions of loci may be to consider them as pertaining to a geometric series of ratio x ($x < 1$). Considering the case of an initial generation, in which selection has not yet developed correlations between loci, the individual contribution of locus ℓ can be written as:

$$B_\ell K_\ell = B_\ell^2 G_{\ell\ell} = \alpha x^{(\ell-1)} \qquad [37]$$

so that the total genetic variance is

$$2\mathbf{B'GB} = 2\alpha \sum_{\ell=1}^{\ell=L} x^{(\ell-1)}$$

$$= 2\alpha \frac{1 - x^L}{1 - x}$$

From equation [17] this yields:

$$L_e = \frac{(1 + x^L)(1 - x)}{(1 - x^L)(1 + x)} \qquad [38]$$

giving quite low values that hardly depend on the actual number L. For example, this equation gives L_e about 3 for $x = 0.5$, L_e about 10 for $x = 0.8$, and L_e about 20 for $x = 0.9$. If an arithmetic series is assumed, the ratio L_e/L remains larger. Writing:

$$B_\ell K_\ell = B_\ell^2 G_{\ell\ell} \propto 1 + \beta\ell \qquad [39]$$

yields

$$\frac{L}{L_e} = 1 + \frac{1}{12} \frac{\beta^2(L^2 - 1)}{\left(1 + \frac{1}{2}\beta(L + 1)\right)^2} \qquad [40]$$

Simulation runs were performed with different actual numbers L of loci sharing the same value of L_e, according to equations [38] and [40] (uniform distribution, arithmetic, and 2 geometric distributions). They result in parallel evolutions concerning the genetic variances V_A (fig 1), as well as for the genic variance V_a and the cumulative response to selection (not shown).

The validity of this parameter as a predictor of the dynamics of the system depends on it being constant over generations. Numerical calculations do not indicate any significant departure of L_e from its initial value. Simulation results show how L_e, as estimated each generation from the simulated data (from equation [17]), changes with time (fig 2). Unless the size of the population is quite small and selection intensity is high, it is seen that L_e is slowly varying; at least, comparison with figure 1 indicates that this parameter changes with time more slowly than the genetic variance. The changes of L_e with time occur after a first period while it is almost constant, in the cases when the initial distribution of contributions of loci to the genetic variance is either uniform (leading to a decrease of L_e) or highly variable (geometric series with ratio $x < 0.8$) in which case L_e increases. Conversely, slightly variable contributions (arithmetic series, or geometric series with x about
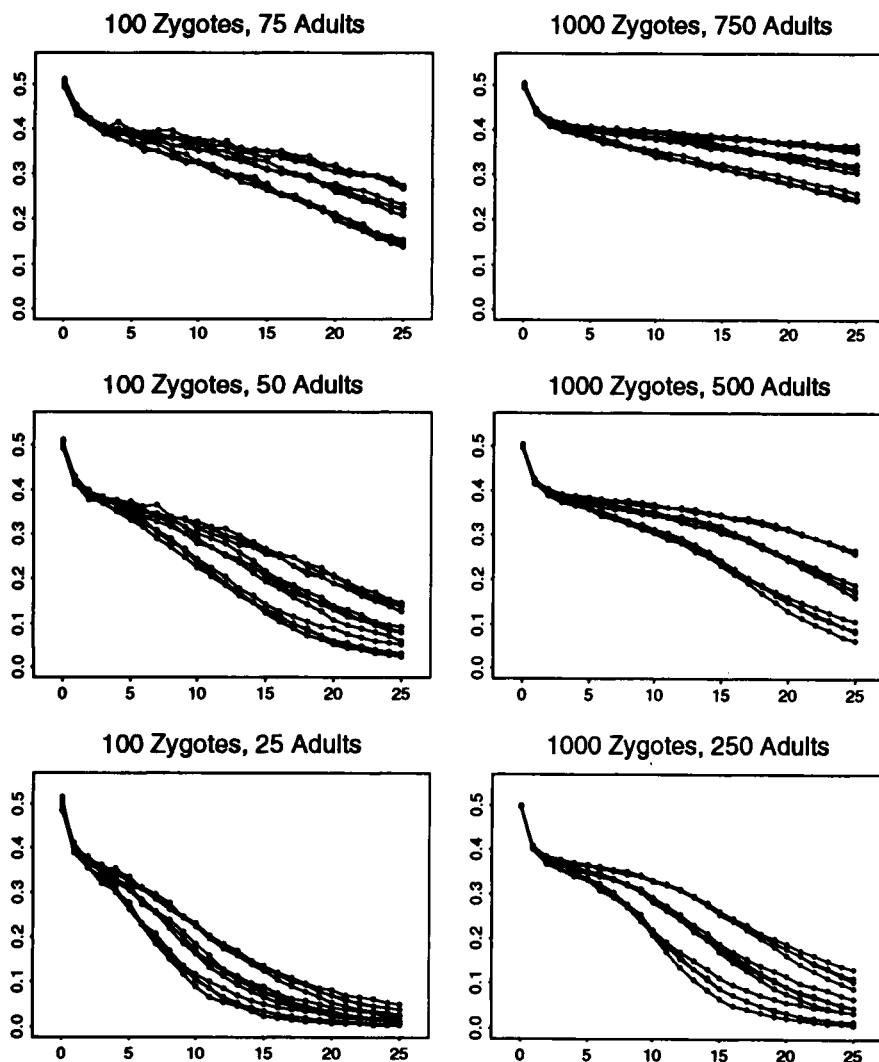
C Chevalet

100 Zygotes, 75 Adults

1000 Zygotes, 750 Adults

100 Zygotes, 50 Adults

1000 Zygotes, 500 Adults

100 Zygotes, 25 Adults

1000 Zygotes, 250 Adults

Fig 1. Change with time under selection of genetic variance for different values of L_e. Initial heritabilities and genetic variances are set to $1/2$. Reproducing individuals (adults) were selected on the basis of individual phenotypes (truncation selection). For each case (population size and selection intensity), values of the genetic variance are plotted for different organizations of the initial genetic variability (mean values in 100 runs); after a few generations, the different curves gather according to the corresponding L_e value. Initial values of L_e and of the total number of loci L are given below. U stands for uniform distribution, AS stands for arithmetic series, GS for geometric series. $L_e = 4$ (down) $L = 4$ (U) $L = 5$ (AS) $L = 5$ (GS) $L = 10$ (GS); $L_e = 8$ (middle) $L = 8$ (U) $L = 10$ (AS) $L = 10$ (GS) $L = 15$ (GS); $L_e = 15$ (up) $L = 15$ (U) $L = 20$ (AS) $L = 20$ (GS) $L = 30$ (GS). Abscissa: time in generations. Ordinate: additive genetic variance.
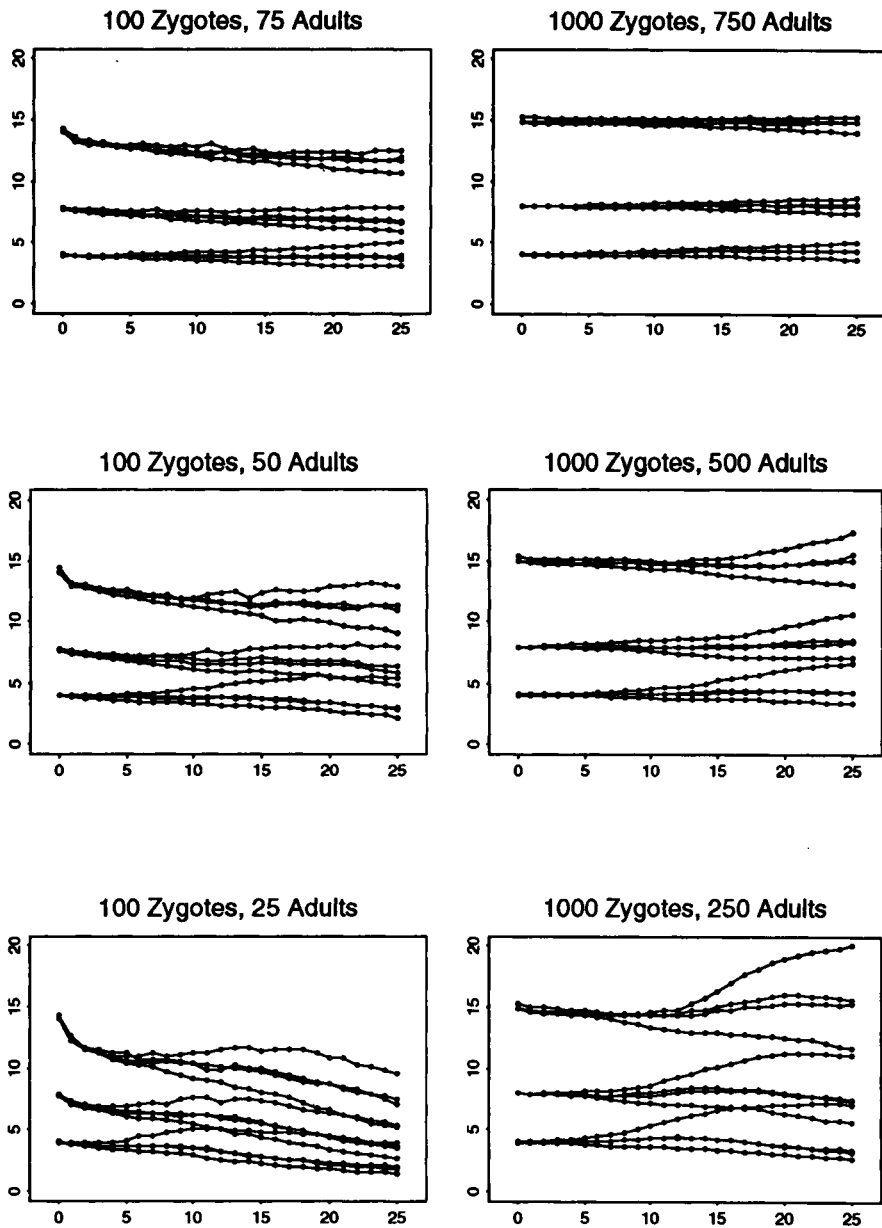
Fig 2. Change with time under selection of the effective number L_e. For the same parameters as in figure 1, current values of L_e, estimated in each run, have been plotted against time (same simulated data).

0.9–0.95) lead to quite stable L_e values. Changes are more significant as selection intensity is greater. Moreover, it seems from simulations that this parameter may be very sensitive to population size, suggesting that a large effective number of QTL's cannot segregate simultaneously in a small population under strong selection.

Thus, the approximate analytical derivations, as well as the simulations, indicate that L_e is a significant parameter. Even if the absolute values obtained for these quantities do not generally fit theoretical predictions very well, the previous results indicate that L_e, as defined in equation [17], controls the sensitivity of the genetic system to the number of QTLs. Moreover, it is a parameter that can be estimated. Firstly, the present L_e definition is the same as that obtained with the classical design using a cross between inbred lines, and comparing the variance in F_2 to the differences of parental lines (Wright, 1968; Lande, 1981). Secondly, methods based on segregation analysis of many linked genetic markers will provide data giving the distribution of the most important QTLs segregating in populations and their contributions to the genetic variance. Even if these results provide new tools for selecting individuals on a true genetic basis, knowledge of L_e remains of interest for analysing performances in populations, as soon as selection has been practised.

Genetic structure under selection

Equation [20] shows that under continuous selection, a genetic structure arises that characterizes an internal equilibrium between selection and recombination. In fact, this relationship holds exactly in the equilibrium state considered by Lande (1976) in the framework of a model involving selection and mutations, when recombination fractions are set to 1/2. It can be also seen as a kind of quasi-linkage equilibrium, a situation encountered in more general models under weak selection (Barton and Turelli, 1991). In contrast with these exact but asymptotic results, the present result holds very early in the process of selection (fig 3 and *Appendix*), then holds during the whole process. Moreover it does not require a weak selection. It is approximate but quite accurate, and probably more meaningful in the context of experimental or applied quantitative genetics than asymptotic results which may be more relevant to evolutionary problems.

This relationship is also dependent on the effective number of QTLs. It is illustrated in figure 3, in which theoretical predictions are compared to the results of simulations. Firstly, it turns out that Gaussian predictions of genetic variances are satisfactory during several generations, and more so as population size is greater and selection intensity lower. It seems however that the theory underestimates the amount of 'hidden' variance in small populations, which is expressed by an estimated value of T larger than its theoretical value. Even with very few QTLs, approximations are good for large population sizes and medium selection intensity, the greatest departures between analytical predictions and simulations appearing for high selection intensities. Secondly, the prediction concerning the organization of genetic variability, as given by equation [20], is more robust than that of genetic variances themselves. The agreement between theory and simulations is observed during more generations, although an increasing variability of the estimated T parameter is observed when significant departures between theoretical and observed variances arise (the estimated variance of T between replicates then shows a sharp increase).
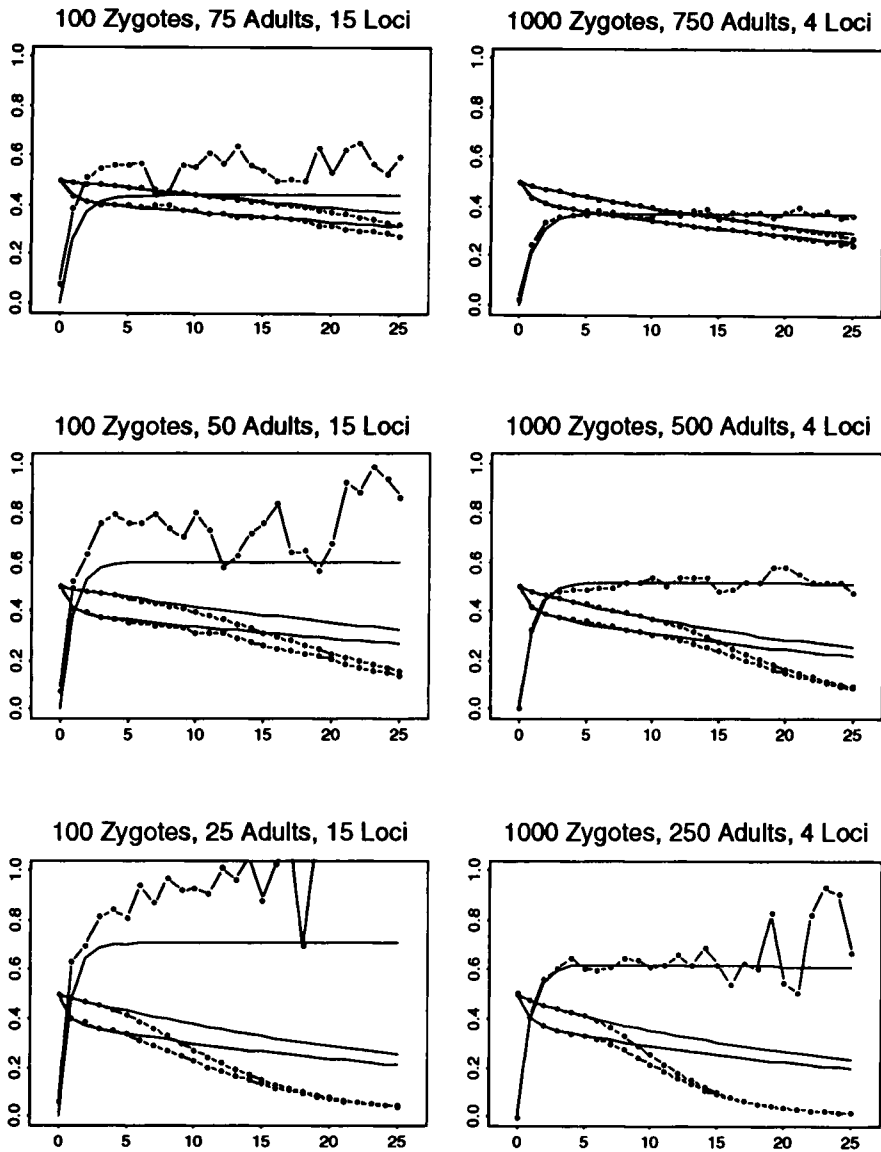
Fig 3. The structure of genetic variance under selection. Ordinates: for each case, 3 values are plotted on the same graph. From up to down (in later generations): the T parameter (equation [20]), the genic variance V_a, and the genetic variance V_A. Theoretical results are shown in continuous lines. Simulation results are given in broken lines. Abscissa: time in generations.

Such a structure of genetic variability under selection implies some changes in the partition of genetic variance among groups of related individuals. For example, within the framework of the simple population model considered until now, the partition of genetic variance into full-sib covariance (C_{FS}), and within-family variance (V_w, which is equal to half the genic variance for unlinked QTLs), can be written as follows, in the population of new unselected zygotes:

$$C_{FS} = \frac{1}{2}V_A \left(1 - \kappa \frac{V_A}{\text{Var}(P)} \right)$$

$$V_w \simeq \frac{1}{2}V_A \left(1 + (1 - \frac{1}{L_e})\kappa \frac{V_A}{\text{Var}(P)} \right)$$

CONCLUSION

In this paper, the genetic model is restricted to a set of unlinked loci. Although it may be stated that any trait is due to many loci spread along the chromosomes, considering a finite set of unlinked loci may not be completely unrealistic. The first experimental results on the localization of QTLs in tomato suggest that a few main regions are involved, and that these regions may in fact include several closely linked genes. The system made up of unlinked loci with many alleles (modelled by a Gaussian distribution of effects) may thus be a correct first-order approximation of a genetic system involving several clusters of genes on different chromosomes, provided no long-term prediction is required. Indeed, simulation runs involving either several unlinked loci with many alleles taken from a normal distribution, or several clusters of tightly linked loci with only 2 alleles, lead to very similar responses to directional selection (not shown). However, this is no longer true in later generations, when many recombination events within clusters generate 'new' alleles, and lead to responses that are unpredictable in the framework of the present theory (Hospital, 1992; Hospital and Chevalet, 1993).

Investigating the effects of the distribution of quantitative loci on the response to selection could be performed here owing to the use of the Gaussian approximation. It is well known that Gaussian distributions are not robust under Mendelian segregation (Felsenstein, 1977), and that, even if a Gaussian distribution is a correct approximation at some time, selection promotes asymmetrical distributions needing higher moments to be included in the recurrence relationships (Turelli and Barton, 1990). When looking at the simulation results, we can see that equation [20] remains correct, but that the reduction of the genic variance V_a under selection is underestimated. This suggests that the Gaussian expression of regressions of genotypes at individual loci on phenotypes is the main incorrect step of the approximation. In fact it is clear that (in a true genetic model) selection shifts allele frequencies and develops asymmetrical distributions which cannot be handled in the framework of Gaussian distributions. Developments of distributions in Gram–Charlier series could be investigated, but such developments would not allow the joint distributions at several loci to be made explicit.

The interesting result obtained here is that several qualitative features of the genetic system are correctly taken into account by the Gaussian approximation. Mainly, the analysis introduces a new macroscopic parameter (the effective number

of quantitative loci) to characterize a quantitative trait, in addition to the usual heritability coefficient. This parameter controls the amount of mid-term selection response and the structure of genetic variance in the population. The other important feature predicted by the model and checked by simulations results is the relationship between the amount of genetic variance available to selection (V_A) and the amount of 'hidden' genetic variance (equation [20]). This result generalizes those obtained by Bulmer (1971) and Verrier et al (1990) for the infinitesimal model, and yields new expressions of covariances between relatives. Whether these discrepancies are important for the methods of estimation of breeding values remains to be investigated.

Other uses of the model may be considered, such as the search for optimal selection intensities in a selection program, or the management of matings in finite populations submitted to strong selection. Indeed, the Gaussian framework would make it easy to take account of population structures (separate sexes, overlapping generations, family or index selection) and of assortative mating. However, the poor long-term predictive power of the method must be stressed. Indeed, the equilibrium genetic structures arising from a balance between mutation and selection depend strongly on the chosen model (Gaussian model *vs* the rare allele or 'house of cards' model, Barton and Turelli, 1987), indicating that the Gaussian approximation should not be used if asymptotic results are searched for. Nevertheless the 'Gaussian' modelling of the optimization of short or mid-term selection schemes may be worthwhile, and provide an improvement over usual theories (Robertson, 1970).

REFERENCES

Barton NH, Turelli M (1987) Adaptive landscapes, genetic distances and the evolution of quantitative characters. *Genet Res Camb* 49, 157-173

Barton NH, Turelli M (1991) Natural and sexual selection on many loci. *Genetics* 127,229-255

Bulmer MG (1971) The effect of selection on genetic variability. *Am Nat* 105,201-211

Chevalet C (1988) Control of genetic drift in selected populations. *In: Proc 2nd Int Conf Quant Genet Raleigh NC* (BS Weir, EJ Eisen, MM Goodman, G Namkoong eds), May 31 – June 5 1987, Sinauer Associates Inc, Sunderland MA, 379-394

Felsenstein J (1977) Multivariate normal genetic models with a finite number of loci. *In: Proc Int Conf Quant Genet* (E Pollack, O Kempthorne, TBJr Bailey eds), Ames IA, August 16–21, 1976, Iowa State University Press, 227-246

Hospital F (1992) Effets de la liaison génique et des effectifs finis sur la variabilité des caractères quantitatifs sous sélection. Thèse, Université des Sciences et Techniques du Languedoc, Montpellier, France

Hospital F, Chevalet C (1993) Effects of population size and linkage on optimal selection intensity. *Theor Appl Genet* 86, 775-780

Lande R (1976) The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet Res* 28,221-235

Lande R (1981) The minimum number of genes contributing to quantitative variation between and within populations. *Genetics* 99,541-553

Robertson A (1970) Some optimal problems in individual selection. *Theor Pop Biol* 1, 120-127

Turelli M, Barton NH (1990) Dynamics of polygenic characters under selection. *Theor Pop Biol* 38,1-57

Verrier E, Colleau JJ, Foulley JL (1990) Predicting cumulated response to directional selection in finite panmictic populations. *Theor Appl Genet* 79,833-840

Wright S (1968) Evolution and the Genetics of Populations. I. Genetic and Biometric Foundations. The University of Chicago Press, Chicago, USA

APPENDIX

Proof of equation [20]

The proof consists of deriving from equations [25] and [26] that T reaches the neighborhood of its limit $\overline{T}$ (equation [27]), within a few generations. The proof is given here in the case of a constant phenotypic variance V_P, but could be extended to the case of a constant environmental variance.

Firstly, equation [26] is rewritten with the new variable

$$U_t = \frac{T^{(t)}}{\overline{T}} - 1 \qquad [41]$$

Then it is shown that the series $|U_t|$ is bounded by a positive series Z_t converging rapidly to a small value.

The following notation is introduced for the current value of heritability

$$h_t^2 = \frac{F(V_a^{(t)}, T^{(t)})}{V_P}$$

Let $\mathcal{D}_t$ be the denominator in equation [26], which can be written as

$$\mathcal{D}_t = 1 - \frac{1}{2N} + \frac{1}{2}\kappa h_t^2 \left(\overline{T}U_t + \overline{T} - (1 - \frac{1}{N}) \right)$$
$$= \frac{1}{2}\kappa h_t^2 \overline{T}U_t + \Delta_t$$

where

$$\Delta_t = 1 - \frac{1}{2N} + \frac{1}{2}\kappa h_t^2 \beta \qquad [42]$$

and

$$\beta = \frac{(1 - \frac{1}{N})(\frac{1}{N} - \frac{1}{L_e} - \frac{1}{2N^2})}{1 - \frac{1}{N} + \frac{1}{2N^2}} \qquad [43]$$

Equation [26] becomes

$$U_{t+1} = \frac{1}{\mathcal{D}_t^2} \left(\frac{1}{2}(1 - \frac{1}{N})U_t + 1 - \frac{1}{N} + \frac{1}{4N^2} \right) - 1$$

$$= \frac{1}{\mathcal{D}_t^2} \left(U_t \left[\frac{1}{2}(1 - \frac{1}{N}) - \kappa h_t^2 \overline{T}\Delta_t - \frac{1}{4}(\kappa h_t^2 \overline{T})^2 U_t \right] \right. \qquad [44]$$

$$\left. + 1 - \frac{1}{N} + \frac{1}{4N^2} - \Delta_t^2 \right) \qquad [45]$$

The expression in the second line [45] can be uniformly bounded by a small number:

$$1 - \frac{1}{N} + \frac{1}{4N^2} - \Delta_t^2 < \alpha = \kappa|\beta| \left(1 - \frac{1}{2N} + \frac{1}{4}\kappa|\beta|\right) \qquad [46]$$

In the first line [44], the sum $\mathcal{S}_t$ of the first 2 terms

$$\mathcal{S}_t = \frac{1}{2}(1 - \frac{1}{N}) - \kappa h_t^2 \overline{T}\Delta_t$$

can be shown to have an absolute value less than $1/2$. The proof depends on the sign of β, and this distinction is also used to get a minoration of $\mathcal{D}_t$.

If β is negative, ie if

$$\frac{1}{L_e} > \frac{1}{N} - \frac{1}{2N^2}$$

then

$$-\frac{1}{L_e} < \beta < 0$$

Introducing these inequalities in Δ, we can write the following inequalities

$$1 - \frac{1}{2N} - \frac{1}{L_e}\kappa h_t^2 < \Delta_t < 1 - \frac{1}{2N}$$

$$\frac{1}{2}(1 - \frac{1}{N}) - \kappa\overline{T}(1 - \frac{1}{2N}) < \mathcal{S}_t < \frac{1}{2}(1 - \frac{1}{N}) - \kappa\overline{T}(1 - \frac{1}{2N} - \frac{1}{L_e}\kappa h_t^2)$$

and taking advantage of

$$\overline{T} < 1 - \frac{1}{L_e}$$

and the hypothesis on the sign of β, we derive

$$-\frac{1}{2} < \mathcal{S}_t < \frac{1}{2}(1 - \frac{1}{N})$$

Similarly, we note that, in this case:

$$\Delta_t > 1 - \frac{1}{2N} - \frac{1}{2}\kappa h_t^2 \frac{1}{L_e}$$

$$\Delta_t > 1 - \frac{1}{2}\left(\frac{1}{N} + \frac{\kappa}{L_e}\right) \qquad [47]$$

In the other case, $\beta > 0$, then

$$0 < \beta < \frac{1}{N}$$

and the preceding inequality [47] still holds. We have

$$1 - \frac{1}{2N} < \Delta_t < 1 - \frac{1}{2N} + \frac{1}{2}\kappa h_t^2 \frac{1}{N}$$

from which we get

$$\frac{1}{2}(1 - \frac{1}{N}) - \kappa h_t^2 \overline{T}\left(1 - \frac{1 - \kappa h_t^2}{2N}\right) < \mathcal{S}_t < \frac{1}{2}(1 - \frac{1}{N}) - \kappa h_t^2 \overline{T}(1 - \frac{1}{2N})$$

$$\frac{1}{2}(1 - \frac{1}{N}) - \kappa h_t^2 \overline{T} < \mathcal{S}_t < \frac{1}{2}(1 - \frac{1}{N})$$

$$\frac{1}{2}(1 - \frac{1}{N}) - \kappa < \mathcal{S}_t < \frac{1}{2}(1 - \frac{1}{N})$$

Thus, the inequality

$$\left|\frac{1}{2}(1 - \frac{1}{N}) - \kappa h_t^2 \overline{T} \Delta_t\right| < \frac{1}{2} \qquad [48]$$

is obtained in both cases, provided that the selection coefficient κ is smaller than $(1 - \frac{1}{2N})$, which is expected in any case.

From equation [47], we can also derive the following lower boundary for the denominator $\mathcal{D}_t$

$$\mathcal{D}_t > 1 - \frac{1}{2}\left(\frac{1}{N} + \frac{\kappa}{L_e}\right) - \frac{1}{2}\kappa |U_t|$$

and following the upper boundary obtained for the numerator, we have

$$|U_{t+1}| < \frac{1}{2} \frac{\alpha + |U_t|\left[1 + \frac{1}{2}\kappa^2 |U_t|\right]}{\left(1 - \frac{1}{2}\left(\frac{1}{N} + \frac{\kappa}{L_e}\right) - \frac{1}{2}\kappa |U_t|\right)^2} \qquad [49]$$

We can define the function φ such that equation [49] is written as

$$|U_{t+1}| < \varphi(|U_t|$$

The function $(Z \to \varphi(Z))$ is positive and monotone increasing for $(0 < Z < 1)$; its value for 0 is about $\alpha/2$. There are 2 roots of the equation $(Z = \varphi(Z))$ in $[0, 1]$; one is about α, and the other can be shown to be larger than $1/2$, provided that κ is not too large (for large N and L_e, the condition is: $\varphi(1/2) < 1/2$ if $\kappa < \sqrt{24} - 4$). Moreover, the derivative of φ at the smallest root is about $1/2$, so that any series (Z_t) with an initial value such that $0 < Z_0 < 1/2$ converges rapidly to this root.

Then it follows from equation [49] that, if for some time v the inequality $|U_v| < 1/2$ holds, the series $(|U_{v+t}|)$ is smaller by the series (Z_t) defined by

$$Z_0 = |U_v|; \quad Z_{t+1} = \varphi(Z_t)$$

and that the series $T^{(t)}$ reaches the neighborhood of $\overline{T}$ within 4–5 generations.

That $|U_t|$ enters the region defined by $|U| < 1/2$ can be checked numerically, by considering any initial conditions $(V_a^{(0)}, T^{(0)})$ in equations [25] and [26], and by twice iterating these equations; in all cases, it is found that $1/2 < T^{(2)} < 3/2$ (for large N and L_e, $T^{(2)}$ can be written as the ratio of 2 polynomials in $x = (T^{(0)} - 1)/2$ and $y = \kappa h_0^2$, and the property can be exactly proven as soon as $-1/2 < x < 1/2$ and $0 < y < 1$).