# Alternative Models for Allozyme-Associated Heterosis in the Marine Bivalve *Spisula ovalis*

**Patrice David,* Bernard Delay,* Patrick Berthou† and Philippe Jarne***

** Laboratoire Génétique et Environnement, Institut des Sciences de l'Evolution, Université Montpellier II, F34095 Montpellier Cedex 05, France, and † I.F.R.E.M.E.R., Centre de Brest, F29273 Brest Cedex, France*

Manuscript received June 7, 1994
Accepted for publication December 14, 1994

ABSTRACT

Correlations between allozyme heterozygosity and fitness-related traits, especially growth, have been documented in natural populations of marine bivalves. However, no consistent pattern has been exhibited, because heterotic effects on size vary with age and individual growth parameters are generally unknown. No consensus has emerged on the genetic basis of allozyme-associated heterosis. The species studied here, *Spisula ovalis*, displays annual shell growth lines, which allows us to compute individual age and growth dynamics over the whole life span. Our morphological study was coupled to a protein electrophoresis study at seven polymorphic loci. While the maximum size gained is not related to heterozygosity, the age at half maximum size, $t_{1/2}$, is significantly negatively correlated with heterozygosity, indicating an heterotic effect on initial growth. The correlation between heterozygosity and size is expected to vanish when age increases, due to the form of the growth function. This decreasing correlation is consistent with previous studies. We compare the relative performances of five linear models to analyze the genetic basis of heterosis. Surprisingly, the largest part of variance in $t_{1/2}$ is due to additive effects, the overdominant components being much weaker. Heterosis is therefore due to general genomic effects rather than to local overdominance restricted to allozymes or small neighboring chromosomal segments. A significant dependence of individual heterotic contributions of the enzyme loci upon expected heterozygosities, rather than metabolic function, further supports the hypothesis of enzymes acting as markers. General genomic effects can hold only if allozyme heterozygosity is positively correlated with heterozygosity at fitness-related genes scattered throughout the genome. This hypothesis is supported here by heterozygosity correlations between enzymatic loci.

ALLOZYME-associated heterosis, that is, the positive correlation between allozyme heterozygosity and fitness-related traits, has been reported in a wide variety of organisms, including pine trees (BUSH *et al.* 1987), vertebrates (PIERCE and MITTON 1982) and especially marine bivalves [*e.g.*, ZOUROS *et al.* (1980) in oysters, KOEHN and GAFFNEY (1984) in mussels and GAFFNEY *et al.* 1990 in surf clams]. Despite this growing dataset, no consensus has emerged regarding the genetic basis of the observed correlations. We do not know whether the allozyme loci under study are directly involved in the observed heterotic effect or whether these loci are mere markers, with other genes being responsible for the measured phenotypic performance. The latter case is referred to as "associative overdominance" after ZOUROS *et al.* (1989). We introduce a further distinction between two kinds of hypotheses for heterosis, namely "local" and "general" effects. The local effect hypothesis assumes that the allozyme loci assayed mark a functionally overdominant narrow chromosomal segment, as defined by HOULE (1989). Heterotic effects may be due to the allozyme loci themselves or to closely linked loci; however, distinguishing between the two possibilities is hardly possible (SMOUSE 1986; HOULE 1989). The general effect hypothesis assumes that allozymes are markers of the genomic heterozygosity as a whole. When increased genomic heterozygosity results in better phenotypic performance (due to either overdominance or dominance over deleterious recessives), a positive correlation between fitness-related traits and allozyme heterozygosity is generated. In other words, allozyme loci reflect inbreeding depression.

Many allozyme studies on heterosis have focused on marine bivalves. Positive correlations between allozyme heterozygosity and shell size, used as a fitness index, have generally been found (ZOUROS *et al.* 1980; DIEHL and KOEHN 1985; GAFFNEY *et al.* 1990). However, the individual age is unknown in most studies. Heterozygotes may then be larger than homozygotes either because they grow older or because of a higher growth rate. This can be overcome by studying individuals settled at a given place and time, so that homogeneity in age is achieved (ZOUROS *et al.* 1980; DIEHL and KOEHN 1985). However, age is known only imprecisely with this technique, because spat collectors are left in the recruiting site for some time and because undesirable

*Corresponding author:* Patrice David, Génétique et Environnement-CC 065, Institut des Sciences de l'Evolution, Université Montpellier II, Place Eugène Bataillon, F34095 Montpellier Cedex 05, France. E-mail: jarne@crit.univ-montp2.fr

settlement after this period cannot be avoided. More importantly, only rough estimators of growth can be obtained, since the growth history of each individual is not accurately summarized by size at a given age. For instance, poor initial growth may be compensated for by enhanced late growth (BERTRAM *et al.* 1993). Age is an important factor, since heterosis for size is frequently detected in juveniles but not in older individuals (DIEHL and KOEHN 1985; MALLET *et al.* 1986; GOSLING 1989). This has been attributed to a better survival of heterozygotes, so that initial differences in heterozygosity between individuals vanish in older age classes, and correlations are subsequently weakened (DIEHL and KOEHN 1985). However, this hypothesis cannot be tested unless individual growth dynamics is known.

The surf clam *Spisula ovalis L.* (Bivalvia, Mactridae) displays external annual shell rings. Winter growth is almost zero and is indicated on the shell as dark, narrow lines. Annuality of this pattern has been demonstrated using an $^{18}O$ isotopic analysis of the shell (BERTHOU *et al.* 1986). Moreover, age-frequency distributions of natural populations based on growth line count are consistent with size-frequency distributions (SMITH 1974; CERRATO 1980). The occurrence of annual growth lines allows computation of individual age and fitting to classical models of growth dynamics. In *S. ovalis*, we can therefore test whether variation in age-independent growth measurements shows heterotic effects and determine which growth parameters are concerned and how aging influences the observed correlations between heterozygosity and size.

We first tested for the heterotic effects using the most classical method, that is, a simple regression of phenotypic performance against multiple locus heterozygosity (MLH). Then we attempted to further characterize the relationship between genotypes and phenotypes by using more complex multilocus models.

## MATERIALS AND METHODS

Individuals of *S. ovalis* were dredged in April 1993, at a single site located in the Bay of Concarneau (Brittany, France). Individuals belonging to only one annual cohort (settled during the 1988 season) were selected, using growth-line counting. A few individuals with unclear growth lines were discarded. Final sample size was 239 individuals.

**Growth analysis:** Each individual shell displayed five growth lines, corresponding to the 1988–1989 to 1992–1993 winters. Morphological measurements included total shell width and width corresponding to the five growth lines (Figure 1). These measurements were performed on right valves, to the nearest 0.1 mm, using an OPTIMAS image analyzer. As growth-line formation occurs in winter and simultaneously for all individuals, shell growth is recorded with a rough 1-yr periodicity. Growth was assumed to follow the classical Von Bertalanffy model, illustrated in Figure 1. The size at time $t$ is therefore

$$L(t) = L_{\max} \cdot [1 - e^{-k \cdot (t - t_o)}], \qquad (A)$$

where $t_o$ is the theoretical birth time, $L_{\max}$ is the maximum size, and $k$ is the initial growth parameter expressed in $(\text{yr})^{-1}$. Knowing $k$, we can calculate the age at which half the maximum size ($L_{\max}/2$) is gained, as $t_{1/2} = \text{Ln}(2)/k$ (Ln stands for natural logarithm). A high $t_{1/2}$ indicates a poor growth. $L_{\max}$ and $t_{1/2}$ were evaluated for each individual using the Ford-Walford regression method (KAUFMANN 1981). This method is based on a linear relation between $L(t+1)$ and $L(t)$ that can be derived from equation (A):

$$L(t+1) = L_{\max} \cdot (1 - e^{-k}) + e^{-k} \cdot L(t). \qquad (B)$$

The measured "width to growth line $i$" was used as a measure of $L(i)$, and we regressed $L(i+1)$ on $L(i)$ for each individual. Equation (B) shows that the regression slope and intercept are estimates of $e^{-k}$ and $L_{\max} \cdot (1 - e^{-k})$, respectively. From these values, we derived estimates of the individual growth parameters $k$ (or $t_{1/2}$) and $L_{\max}$.

$L_{\max}$ is known to be under environmental effects and is correlated with $t_{1/2}$ in our sample ($r = 0.71$, $P < 0.001$), reflecting ecological or developmental constraints. To reduce environmental noise, we used the residuals of the regression, noted $t^*_{1/2}$, as corrected values of $t_{1/2}$.

**Genetic analysis:** Individuals were frozen alive and kept at −70° before electrophoresis. Adductor muscles, part of the foot and digestive gland were dissected and homogenized in an equal volume of homogenization buffer (0.01 M Tris-HCl, 0.001 M EDTA, 0.04 g/liter NADP, pH 6.8). Genotypes for seven polymorphic loci were determined using horizontal starch-gel electrophoresis. Glucose phosphate isomerase (EC 5.3.1.9) and phosphoglucomutase (EC 5.4.2.2) were resolved on Tris-citrate pH 8.0 gels, using muscle and foot extracts. Isocitrate dehydrogenase (EC 1.1.1.42), aspartate aminotransferase (EC 2.6.1.1), and leucine aminopeptidase (*LAP*; EC 3.4.11.−) were run on TC 8.0 buffers using digestive gland extracts. Beta-galactosidase (*GAL*; EC 3.2.1.23) was resolved on Tris-maleate-EDTA pH 6.9 buffer, using digestive gland extracts. Detailed recipes of gel buffers and staining solutions are given in PASTEUR *et al.* (1988).

**Genetic data analysis:** Classical population genetics parameters were computed, including allele frequencies, expected heterozygosity and estimated $F_{is}$. Expected heterozygosity $H_e$ was computed using Levene's correction (NEI 1987, p. 152) for finite samples and $F_{is}$ was estimated as the minimum-bias estimator $f$ derived by WEIR and COCKERHAM (1984). Consistency of observed heterozygote proportions with their expected values under Hardy-Weinberg equilibrium was tested comparing the statistic $X^2 = nf^2$ with critical values of the chi-square distribution with one degree of freedom.

**Models for analysis of heterotic effects:** The simplest model is a univariate linear regression of phenotypic performance (say growth) on MLH. Hence, the fitness trait $Y$ is expressed as

$$Y = \alpha \cdot (X_1 + X_2 + \cdots + X_n) + \beta + \epsilon, \qquad (1)$$

where $X_i$ denotes heterozygosity at the $i$th locus (scoring 1 for heterozygote, 0 for homozygotes), $\alpha$ and $\beta$ are the parameters of the model, and $\epsilon$ the error term. This model fits the general effect hypothesis, as only the number of heterozygous loci is taken into account, regardless of loci and genotype. Many studies of heterosis in marine bivalves have used this univariate model (GREEN *et al.* 1983; DIEHL and KOEHN 1985; ZOUROS *et al.* 1988; GAFFNEY 1990), although no consensus emerged for either the general or local hypothesis.

A more sophisticated model was developed by SMOUSE (1986) with the explicit purpose of fitting multiplicative overdominance at independent loci. Assuming a locus with two alleles A and B maintained by overdominance, the expected

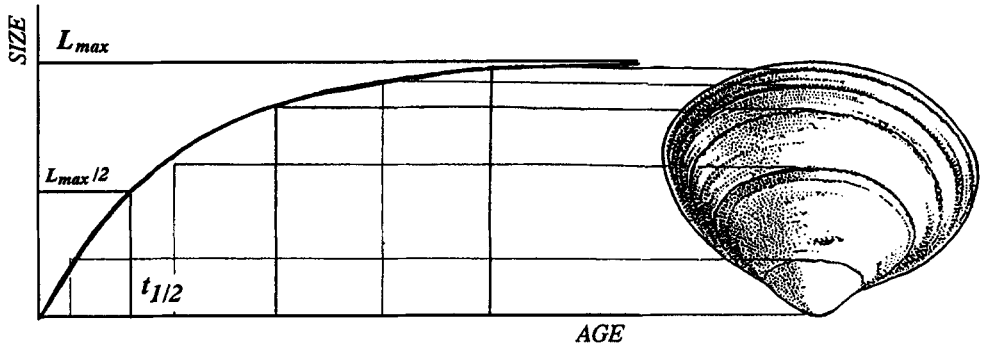


FIGURE 1.—Fitting the Von Bertalanffy growth curve with individuals of *Spisula ovalis*. Lines on the shell are separated by 1 year of growth and values in abscissa correspond to successive winter growth pauses. The maximum size $L_{max}$ is asymptotically reached at the end of the life. The speed of this asymptotic convergence is inversely related to $t_{1/2}$, the age at half the maximum size. The method used to estimate $L_{max}$ and $t_{1/2}$ is described in the text.

relationships between allele frequencies at equilibrium ($p$ and $q$, respectively) and genotypic fitnesses can be derived. Genotypes are characterized by their adaptive distances (scoring $D = 0$ for AB, $1/p$ for AA and $1/q$ for BB). Assuming overdominance at each locus, multiplicative fitnesses between loci and panmixia, a linear negative relation is expected at equilibrium between the natural logarithm of fitness and adaptive distances:

$$\mathrm{Ln}(Y) = \alpha_1 \cdot D_1 + \alpha_2 \cdot D_2 + \cdots + \alpha_n \cdot D_n + \beta + \epsilon, \quad (2)$$

where $D_i$ is the adaptive distance at the $i$th locus and the $\alpha_i$ and $\beta$ are the parameters of the model. For multiallelic loci, A is chosen as the most common allele, and all remaining alleles are pooled in a composite B class, as recomended by SMOUSE (1986). The allele frequencies $p$ and $q$, needed for computation of the $D_i$, are estimated from the sample. Model (2) takes into account $n$ independent variables *vs.* one for model (1) and is therefore expected to explain more variance, even though this additional variance is mere random sample noise. Model (1) is more constrained than (2) for two reasons. First, all loci have equivalent effects in (1), so there is no possibility for locus-specific effects. Second, each locus in (1) is characterized as either homozygote or heterozygote, whereas (2) discriminates, in the diallelic case, between the two homozygote genotypes. These differences may be overcome by fitting "intermediate" models.

The multivariate version of model (1) is

$$Y = \alpha_1 \cdot X_1 + \alpha_2 \cdot X_2 + \cdots + \alpha_n \cdot X_n + \beta + \epsilon. \quad (3)$$

This model involves as many independent variables as model (2) and accounts for locus-specific effects. This allows comparison between the results of (2) and (3), as already done by BUSH *et al.* (1987).

To provide a "null model" for the overdominance hypothesis, a fourth model was developed. As in model (2), this discriminates between the three genotypes (AA, AB and BB) in the diallelic case but ranks them in a strictly additive manner:

$$Y = \alpha_1 \cdot R_1 + \alpha_2 \cdot R_2 + \cdots + \alpha_n \cdot R_n + \beta + \epsilon, \quad (4)$$

where $R_i$ is 2 for AA, 1 for AB, 0 for BB, at locus $i$. Given the symmetry of the $R_i$ distribution, the choice of the A allele influences the value of the $\alpha_i$ coefficients, but the total percentage of explained variance remains unchanged. A and B are defined as in model (2).

Models (2) and (4) can be combined to evaluate the relative importance of additive *vs.* overdominant effects:

$$Y = \alpha_1 \cdot R_1 + \beta_1 \cdot D_1 + \cdots + \alpha_n \cdot R_n + \beta_n \cdot D_n + \gamma + \epsilon. \quad (5)$$

In this model, the genotypic effects on fitness, at locus $i$, are given by $Y(AA) = 2\alpha_i + \beta_i/p$, $Y(AB) = \alpha_i$ and $Y(BB) = \beta_i/q$. To compare, for each locus, additive and overdominant effects, the following crude procedure was used: we standardized these values to $Y^*(AA) = 1$ and $Y^*(BB) = 0$ by a linear transformation (* denotes standardized values), which gives $Y^*(AB) = (\alpha_i - \beta_i/q) \cdot [2\alpha_i + \beta_i (1/p - 1/q)]^{-1}$. The standardized value for AB expected with additive effects only is $R^*(AB) = {}^1/_2$, and with only overdominant effects $D^*(AB) = p/(p - q)$. The weight of overdominant effects relative to additive ones can be estimated by the ratio $O_i = |[D^*(AB) - Y^*(AB)]/[R^*(AB) - Y^*(AB)]|$. When $O_i < 1$, overdominant effects dominate additive ones.

**Application to the dataset:** Model (1) was applied to parameters $t_{1/2}$, $t^*_{1/2}$ and $L_{max}$. The unbiased estimates of standard deviation of growth parameters within each heterozygosity class were also computed and regressed on MLH values.

To estimate the variation of heterozygosity-size correlations with age, the effect of MLH on size was estimated at given ages. Estimates of size were obtained using the growth parameters $L_{max}$ and $k$ for each individual and setting age $x = (t - t_o)$ to various values in equation (B).

The multivariate regressions (2) to (5) were applied to $t^*_{1/2}$. The relative performances of the five models were compared on the basis of percentage of explained variance. For all regressions, significance was evaluated through an $F$ test (SOKAL and ROHLF 1981, pp. 631–642). For univariate regressions, the correlation coefficient was given for comparison with previous studies. Pooling the alleles involves a loss of information for models (2), (4) and (5) compared with model (3). This effect was evaluated by applying model (3) to the same AA, AB and BB data as for other models, that is, considering heterozygotes for two rare alleles as homozygotes (the resulting regression is noted 3′).

The weights attributed to each locus in model (3) were regressed on $f$ and on expected heterozygosity to estimate the influence of genetic parameters on locus-specific effects.

## RESULTS

**Growth analysis:** The mean correlation coefficient of individual Ford-Walford type regressions is 0.975 ± 0.02 (± SD), indicating that the growth of *S. ovalis* fits the Von Bertalanffy equation and that $L_{max}$ and $t_{1/2}$ correctly describe individual growth history. A large variation among individuals is observed: $L_{max}$ ranges from 2.18 to 4.63 cm (mean 3.33 cm) and $t_{1/2}$ from 0.83 to

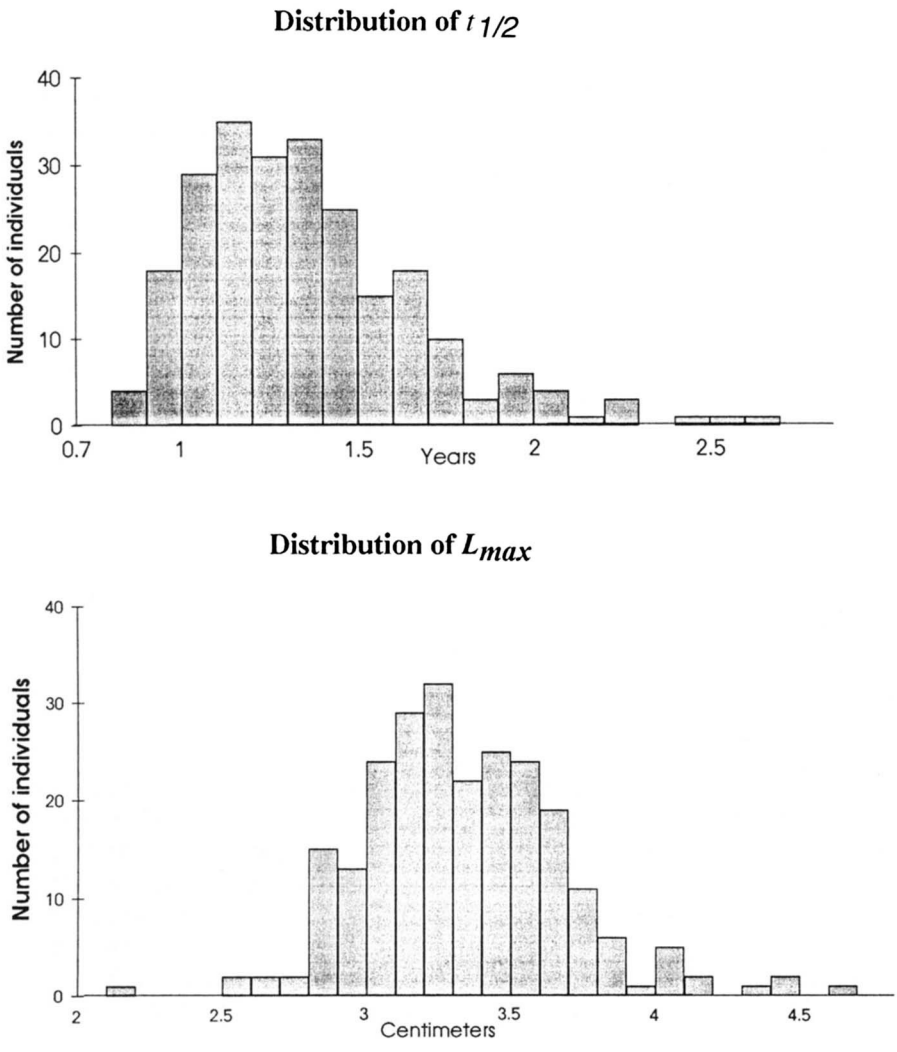


FIGURE 2.—Distribution of the growth parameters in the sample studied. Total sample size is 239 individuals.

2.78 yr (mean 1.36 yr). The distributions are presented in Figure 2. As previously mentioned, $t_{1/2}$ and $L_{max}$ are significantly positively correlated.

**Genetic analysis:** Genetic parameters are given in Table 1. All loci are highly polymorphic, with expected heterozygosities ranging from 0.123 to 0.713. Most $F_{is}$ values were positive, but only one is significant (*GAL*; Table 1).

**Growth and heterozygosity, univariate regressions (model 1):** The effect of MLH on growth parameters is presented in Figure 3. $t_{1/2}$ is significantly correlated with MLH ($r = -0.17$; $F_{1,237} = 6.75$; $P = 0.01$), and its variance decreases with increasing MLH ($r = -0.79$; $F_{1,5} = 8.21$; $P = 0.035$). $L_{max}$ is not correlated with MLH ($r = -0.05$; NS); neither does its variance covary with MLH ($r = 0.14$; NS). The correlation between $t_{1/2}$ and $L_{max}$ introduces a bias in the evaluation of heterotic effects on $t_{1/2}$. When the effect of $L_{max}$ is removed, the negative correlation between MLH and $t^*_{1/2}$ is increased ($r = -0.19$; $F_{1,237} = 8.58$; $P = 0.004$).

The changing effect of MLH on mean estimated body size with increasing age is presented in Figure 4. When age $x$ is near 0, the Von Bertalanffy equation reduces to the linear form $L(x) = L_{max} \cdot k \cdot x$, so that the effect of MLH on mean body size at age $x$ is the same as on ($L_{max} \cdot k$). When $x$ increases, this effect progressively vanishes, and finally for large $x$, $L(x)$ tends toward $L_{max}$ so that the effect of MLH on $L(x)$ approaches that on $L_{max}$, which was previously shown to be near zero.

**Multiple regressions [models (2)–(5)]:** Models (2)–(5) were tested on $t^*_{1/2}$. The results are summarized in Table 2, and locus-specific coefficients are given in Table 1. Model (2) explains 5.4% of the total variance in $t^*_{1/2}$ (not significant), that is, slightly more than model (1) (3.5%). Heterotic effects are expected to yield positive coefficients in model (2), which is true for five loci out of seven (Table 1), one of them being significantly different from zero. As some $t^*_{1/2}$ were negative, log transformation was not performed. This makes little difference, as this model assumes that fitness differences between individuals are reasonably low (SMOUSE 1986), so that fitness and its natural logarithm are approximately linearly related. Model (3) explains 6.7% of the variance in $t^*_{1/2}$ (significant). In this model, heterosis is expected to yield negative coefficients, which is true for six loci, one of them being significantly

**TABLE 1**

**Locus-specific parameters for the sample studied**

| Locus | No. of alleles | Expected heterozygosity | $f^b$ | $X^2 = nf^2$ | Model (2)[a] | Model (3)[a] | Model (4)[a] | Model (5)[a] | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | $\alpha_i$[c] | $\beta_i$[c] | $O_i$[c] |
| *GPI* | 4 | 0.123 | 0.049 | 0.58 | $2.2.\ 10^{-2}$ | $3.1.\ 10^{-2}$ | $-7.8.\ 10^{-2}$ | $-7.8.\ 10^{-2}$ | $1.6.\ 10^{-2}$ | 0.68 |
| *PGM* | 8 | 0.244 | 0.108 | 2.77 | $-0.9.\ 10^{-2}$ | $-2.6.\ 10^{-2}$ | $3.8.\ 10^{-2}$ | $1.3.\ 10^{-2}$ | $-0.7.\ 10^{-2}$ | 0.63 |
| *LAP* | 10 | 0.612 | 0.069 | 1.15 | $0.2.\ 10^{-2}$ | $-4.1.\ 10^{-2}$ | $7.3.\ 10^{-2}$* | $7.5.\ 10^{-2}$* | $-0.6.\ 10^{-2}$ | 46.4 |
| *AAT* | 4 | 0.317 | 0.038 | 0.35 | $-0.1.\ 10^{-2}$ | $-3.0.\ 10^{-2}$ | $3.1.\ 10^{-2}$ | $3.4.\ 10^{-2}$ | $-0.2.\ 10^{-2}$ | 9.99 |
| *IDH* | 5 | 0.472 | −0.083 | 1.65 | $0.2.\ 10^{-2}$ | $-0.0.\ 10^{-2}$ | $0.1.\ 10^{-2}$ | $-0.8.\ 10^{-2}$ | $-0.1.\ 10^{-2}$ | 11.0 |
| *EST* | 2 | 0.454 | −0.066 | 1.05 | $3.4.\ 10^{-2}$* | $-5.6.\ 10^{-2}$ | $-2.2.\ 10^{-2}$ | $-2.7.\ 10^{-2}$ | $3.1.\ 10^{-2}$* | 1.28 |
| *β-GAL* | 10 | 0.712 | 0.192 | 8.78** | $0.0.\ 10^{-2}$ | $-9.7.\ 10^{-2}$* | $3.1.\ 10^{-2}$ | $2.6.\ 10^{-2}$ | $2.7.\ 10^{-2}$ | 6.65 |

*,** Significance at $P = 0.05$ and $P = 0.01$, respectively.

[a] Effects of each locus on corrected $t_{1/2}$ expressed by its coefficients in the multiple-regression models (2)–(4).

[b] Unbiased estimator of $F_{is}$ from WEIR and COCKERHAM (1984).

[c] For definition of the coefficients $\alpha_i$, $\beta_i$ and $O_i$ in model (5), see text.

different from zero (*GAL*). These coefficients are significantly correlated with expected heterozygosity but not with $F_{is}$ at corresponding loci ($r = -0.79$, $P = 0.03$ and $r = -0.45$, not significant, respectively). When all alleles except the most common are pooled in a single class [as in models (2), (4) and (5)], the result (model 3′) is no longer significant. Model (4) explains 7.9% of the variance in $t^*_{1/2}$ (highly significant) and one coefficient is significantly different from zero (*LAP*). Model (5) shows a highly significant dependence of $t^*_{1/2}$ on multilocus genotypes (12.5% variance explained). Two out of the seven $\alpha_i$ coefficients (additive locus-specific effects) are significant. Only three $\beta_i$ coefficients are positive (*i.e.*, in agreement with expectations of the overdominance hypothesis), one being significant. For five out of seven loci, the ratio $O_i$ is higher than 1, indicating that observed genotypic scores are closer to additive than to overdominant expectations at those loci.

## DISCUSSION

**Effects of MLH on growth and size:** $t_{1/2}$, after removal of the effect of $L_{max}$, is negatively correlated with heterozygosity over the seven enzymatic loci studied in our sample of Spisula. $t_{1/2}$ being an inverse indicator of growth, this correlation indicates an heterotic effect in wild populations of *S. ovalis*, which is consistent with heterozygosity-size data reported in various bivalve species (ZOUROS *et al.* 1980; KOEHN and GAFFNEY 1984;

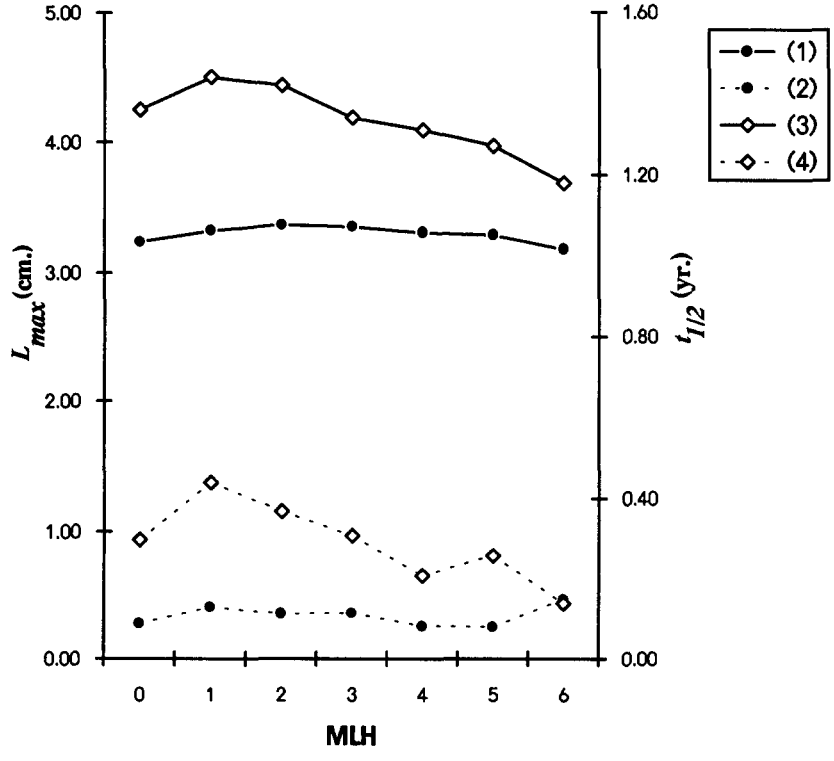


FIGURE 3.—Effect of multiple-locus heterozygosity (MLH) on mean and variance of growth parameters. (1) mean $L_{max}$, (2) standard error of $L_{max}$, (3) mean $t_{1/2}$, (4) standard error of $t_{1/2}$. Sample sizes of MLH classes 0–6 are 6, 34, 57, 81, 36, 20 and 5, respectively.

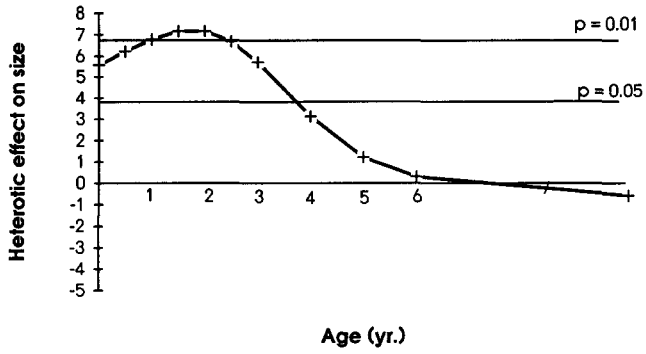


FIGURE 4.—Heterotic effect on mean estimated sizes as a function of age. The heterotic effect is computed as $F_{1,237}$, from the ANOVA of the regression of estimated size at a given age against MLH. The sign given to $F_{1,237}$ is that of the correlation between MLH and mean size. Horizontal lines correspond to significance levels. The limit values of $F_{1,237}$ for ages 0 and $\infty$ were calculated from regressions of $(k \cdot L_{max})$ and $L_{max}$, respectively, against MLH (see text).

**TABLE 2**

**ANOVA for regression models (1)–(5) of $t^*_{1/2}$ against genetic variables**

| Model | No. of explicative variables | Explained sum of squares | $F$ | $P$ |
|---|---|---|---|---|
| (1) | 1 | 0.455 (3.5%) | 8.58 | 0.004 |
| (2) | 7 | 0.708 (5.4%) | 1.89 | 0.070 |
| (3) | 7 | 0.872 (6.7%) | 2.37 | 0.023 |
| (3′)[a] | 7 | 0.396 (3.0%) | 1.04 | 0.408 |
| (4) | 7 | 1.035 (7.9%) | 2.85 | 0.007 |
| (5) | 14 | 1.626 (12.4%) | 2.28 | 0.006 |

Percentages are relative to the total sum of squares (=13.03 for all models). Degrees of freedom for $F$ are 1 and 237 for model (1), 7 and 231 for models (2)–(4), 14 and 224 for model (5). $P$ denotes the probability of Type 1 error associated with rejection of the null hypothesis of no effect.

[a] Model (3′) is similar to model (3), with all alleles but the most common one pooled in a single class at every locus.

KOEHN *et al.* 1988). The growth analysis used in the present study rules out confusing effects of genotype-dependent survival without the need to keep individuals in isolation, so that the effect studied is under fully natural conditions, including for instance density-dependent competition among and within cohorts. This conclusion is not restricted to a given age, as in previous studies, because the individual growth dynamics is characterized over the whole life span. Moreover, the parameter $t_{1/2}$ shows a decreasing variance with increasing heterozygosity. This last trend is a classical feature of growth traits in bivalves, for which possible explanations include increased developmental homeostasis for heterozygotes (KAT 1982; MITTON and GRANT 1984) or mere combinatory and additive effects (CHAKRABORTY and RYMAN 1983).

The growth model used allows us to compute the correlations between heterozygosity and size at various ages. The correlation decreases with age, being significantly positive during the first 2 years and approaching zero after 3 years. This age effect has already been detected in bivalves, as heterosis was reported only in young cohorts (MALLET *et al.* 1986; GOSLING 1989). Given the observed relationship between heterozygosity and survival (ZOUROS *et al.* 1983), it was argued that a decline of the variance in heterozygosity was the cause of weakened heterozygosity-size correlations in old cohorts (DIEHL and KOEHN 1985). However, in Spisula, such age-dependent correlations were observed for a fixed set of individuals, therefore excluding effects of differential survival based on genotype. The age effect originates in the growth dynamics: size is initially a decreasing function of the parameter $t_{1/2}$, which is related to MLH. Later in life, size becomes independent of $t_{1/2}$ and depends only upon $L_{max}$, which is not related to MLH. Given that the growth of many bivalve species, including Mytilus spp (SEED 1980), seems to conform to the Von Bertalanffy growth model, the effect observed here may be quite general. The lack of a precise and consistent representation of the individual growth phenotype is therefore partly responsible for previous studies on marine bivalves yielding inconsistent results. Size of a young animal is not the same phenotype as size in an old one.

**Relative performances of alternative models for heterosis:** We used the results of models (1)–(5) to address the following questions. Is it possible to distinguish genotype-specific effects on growth? Do these effects follow the overdominant expectations? Is heterosis due to local or general effects in our sample?

The highly significant result of model (1) ($P$ = 0.004) clearly indicates that the phenotype studied is partly genetically determined. Allowing for locus-specific effects [model (3)] increases the variance explained but decreases the significance ($P$ = 0.02), suggesting that an important part of the extra variance explained is probably sampling error. On the other hand, genotypes can be described not only as heterozygote *vs.* homozygote but as three classes: AA, AB and BB. This is done in model (5), which allows for locus-specific effects plus genotype-specific effects, with no constraint on the relative effects of the three genotypes. Model (5) is indeed highly significant ($P$ = 0.006) despite the large number of variables. Genetic variation of the growth parameter may therefore result mainly from differences between genotypes within loci rather than from distribution of heterozygosity among loci.

The next question is whether these genotype-specific effects take an overdominant form. The adaptive distance model (2) explains more variance than model (1) but is not significant ($P$ = 0.07), so one cannot exclude the possibility that the variance explained by (2) is only sampling error. More importantly, the fraction of the variance explained is less than that explained

by model (3), meaning that ranking three genotypes in an overdominant manner provides less reliable information than considering heterozygote and homozygote classes only at each locus. Indeed, the strictly additive ranking [model (4)], which is the zero-overdominance situation, leads to increased explained variance and high significance ($P = 0.007$). Similarly, in model (5), the observed $O_i$ ratios are generally higher than 1, meaning that additive effects are more important than overdominant ones. If we take overdominance as the null hypothesis, the observed standardized fitnesses of heterozygotes ($Y_{AB^*}$) are expected to be symmetrically distributed around the overdominant expectation ($D_{AB^*}$). For all seven loci, $Y_{AB^*}$ is biased toward the additive expectation (two-tailed sign test, $P = 0.016$), again leading to rejection of the purely overdominant model. However, results from model (5) should be taken cautiously, as 15 parameters are estimated from only 239 data.

**Possible biases:** Several possible biases could reduce the performance of the adaptive distance model in this study. The first one is mortality. Our sample is composed of old individuals compared with other bivalve studies (*e.g.*, KOEHN *et al.* 1988; ZOUROS *et al.* 1988; GAFFNEY *et al.* 1990). If growth and mortality before sampling are correlated, only the fastest growing individuals survive in each genotypic class, and differences between estimated genotypic scores are smoothed. This bias is the cost paid for a precise representation of growth, which requires a minimum of three to four annual growth lines. A more important source of error for multivariate models is allele pooling, which cannot be avoided for some multilocus models without very large sample sizes. The significant heterotic effect shown by model (3) all but disappears when rare alleles are pooled in a single class (model 3′). However, presampling mortality or allele pooling can reduce the apparent fitness differentials between genotypes but are not expected to modify their respective fitness ranks. The estimated overdominant effect may be driven below the significance level (as observed in our dataset), but the biases cannot generate highly significant additive effects. Therefore the overdominant model does not explain the heterotic effects observed here.

**Genetic bases of heterosis in Spisula: local *vs.* general effects:** The failure of the adaptive distance model leaves us with the paradox of a dataset clearly displaying heterotic effects but with no hint of apparent or real overdominance. This is surprising, since the ranking of fitnesses given by the adaptive distances is predicted under a variety of models, based on either local or general effects (SMOUSE 1986; ZOUROS 1993; HOULE 1995). This leaves us with two hypotheses. First, heterosis is due to local effects (but not to overdominant ones), that is, heterozygotes are generally intermediate between homozygotes and slightly better than the average homozygote. Second, the enzyme loci studied only reflect heterozygosity at other genes scattered throughout the genome, through some sort of genotypic correlations.

The first hypothesis does not explain how genetic polymorphism is maintained. This can be done through different genotypic fitnesses in heterogeneous environments as suggested by MITTON and GRANT (1984) and GILLESPIE (1991). However, such effects are unlikely to operate at each of several independent enzymatic loci. The *LAP* locus in Mytilus (KOEHN *et al.* 1980) is a good, but isolated, example among marine bivalves. The possibility of direct selection on allozymes was sustained by KOEHN *et al.* (1988), the effect of each locus being related to the metabolic function of the corresponding enzyme in *Mulinia lateralis*. However, in the present study, no trend is apparent concerning enzymatic function. The main determinant of locus-specific effects seems to be expected heterozygosity.

The general effect hypothesis does not face the problem of loss of polymorphism, as enzyme loci are expected to behave as mere neutral markers. Hypotheses relying on general effects include inbreeding (CHARLESWORTH 1991) and linkage disequilibrium with deleterious recessives (OHTA 1971). HOULE (1995) has shown that under both processes, the adaptive distance model is expected to fit the data well, which is not the case here. However, this does not preclude general effects, as (1) HOULE's analysis was restricted to one locus and (2) general effects can be due to a variety of distinct mechanisms. For instance, association with genes partially (but not fully) dominant for fitness can generate marker-associated heterosis, without apparent overdominance at the marker loci (ZOUROS 1993).

The main underlying assumption of the general effect hypothesis is the existence of genotypic correlations in the genome, which seems at variance with the panmictic appearance of reproductive systems in marine bivalves (HOULE 1989). However, several observations, including frequent (and unexplained) heterozygote deficiencies (review in ZOUROS and FOLTZ 1984), suggest nonrandom mating. These deficiencies consistently co-occur with heterosis. For instance, in *M. lateralis*, the heterotic effect on size is significantly greater for loci showing higher heterozygote deficiency (GAFFNEY *et al.* 1990). In the present study, estimates of $F_{is}$ are mostly not significant but also rather imprecise (due to large number of alleles and possible presampling mortality). Another indication of deviation from panmixia is the occurrence of genetic correlations between allozymes in the samples showing heterosis. In the present dataset, 15 of 21 covariances between pairs of heterozygosity indexes $X_i$ were positive. A similar trend was observed in *M. lateralis*, in which very homozygous and very heterozygous genotypes were in excess compared with the distribution expected under independence (GAFFNEY *et al.* 1990). There is no obvious reason for such heterozygosity correlations to occur only within

the set of allozyme loci analyzed. Hence, heterozygosity at allozyme loci may be positively correlated with heterozygosity at fitness genes. These correlations seem very weak compared with those produced through inbreeding. However, the heterozygote advantage in our study is only in comparison with the mean homozygote phenotype and does not produce apparent overdominance (superiority over both homozygotes). The genetic correlations necessary for the observed effect need not be as strong as those generated by high selfing rates (described by CHARLESWORTH 1991). How such correlations could originate remains unclear. An interesting clue, in benthic sedentary species with planktonic phase, could be found in the settlement dynamics, with genetically different groups of individuals recruited at the same place (*cf. e.g.*, JOHNSON and BLACK 1984). However, the magnitude of expected correlations has still to be investigated both empirically and theoretically.

We thank P. NOEL, A. LE MENACH, G. PISTRE and A. F. PERNOT for technical support; D. HOULE for access to an unpublished manuscript; F. BONHOMME, J. BRITTON-DAVIDIAN, D. CHARLESWORTH, M. RAYMOND, F. RENAUD, F. ROUSSET, L. THALER and especially two anonymous reviewers for insightful discussions and/or helpful comments on the manuscript. This work was supported by I.F.R.E.M.E.R. and C.N.R.S. (URA 327) grants. This is contribution no. 95-003 of Institut des Sciences de l'Evolution.

## LITERATURE CITED

BERTHOU, P., M. BLANCHARD, P. NOEL and C. VERGNAUD-GRAZZINI, 1986 The analysis of stable isotopes of the shell applied to the determination of the age of four bivalves of the "Normano-Breton" gulf, western Channel. International Council for the Exploration of the Sea-Shellfish commitee, Annual report (K:16): 1–13.

BERTRAM, D. F., R. C. CHAMBERS and W. C. LEGGET, 1993 Negative correlations between larval and juvenile growth rates in winter flounder: implications of compensatory growth for variation in size-at-age. Mar. Ecol. Prog. Ser. **96:** 209–215.

BUSH, R. M., P. E. SMOUSE and F. T. LEDIG, 1987 The fitness consequences of multiple locus heterozygosity: the relationship between heterozygosity and growth rate in pitch pine (*Pinus rigida* Mill.). Evolution **41**: 787–798.

CERRATO, R. M., 1980 Demographic analysis of bivalve populations, pp. 417–465 in *Skeletal Growth of Aquatic Organisms—Biological Records of Environmental Change*, edited by D. C. RHOADS and R. A. LUTZ. Plenum Press, New York.

CHAKRABORTY, R., and N. RYMAN, 1983 Relationship of mean and variance of genotypic values with heterozygosity per individual in a natural population. Genetics **103**: 149–152.

CHARLESWORTH, D., 1991 The apparent selection on neutral marker loci in partially inbreeding populations. Genet. Res. **57**: 159–175.

DIEHL, W. J., and R. K. KOEHN, 1985 Multiple-locus heterozygosity, mortality and growth in a cohort of *Mytilus edulis*. Mar. Biol. **88**: 265–271.

GAFFNEY, P. M., 1990 Enzyme heterozygosity, growth rate, and viability in *Mytilus edulis*: another look. Evolution **44**: 204–210.

GAFFNEY, P. M., T. M. SCOTT, R. K. KOEHN and W. J. DIEHL, 1990 Interrelationships of heterozygosity, growth rate and heterozygote deficiencies in the coot clam, *Mulinia lateralis*. Genetics **124**: 687–699.

GILLESPIE, J. H., 1991 *The Causes of Molecular Evolution*. Oxford University Press, New York.

GOSLING, E. M., 1989 Genetic heterozygosity and growth rate in a cohort of *Mytilus edulis* from the Irish coast. Mar. Biol. **100**: 211–215.

GREEN, R. H., S. M. SINGH, B. HICKS and J. MCCUAIG, 1983 An arctic intertidal population of *Macoma balthica* (Mollusca, Pelecypoda): genotypic and phenotypic components of population structure. Can. J. Fish. Aquat. Sci. **40**: 1360–1371.

HOULE, D., 1989 Allozyme-associated heterosis in *Drosophila melanogaster*. Genetics **123**: 789–801.

HOULE, D., 1995 Adaptive distance and the genetic basis of heterosis. Evolution (in press).

JOHNSON, M. S., and R. BLACK, 1984 Pattern beneath the chaos: the effect of recruitment on genetic patchiness in an intertidal limpet. Evolution **38**: 1371–1383.

KAT, P., 1982 The relationship between heterozygosity for enzyme loci and developmental homeostasis in peripheral populations of aquatic bivalves (Unionidae). Am. Nat. **119**: 824–832.

KAUFMANN, K. W., 1981 Fitting and using growth curves. Oecologia **49**: 293–299.

KOEHN, R. K., and P. M. GAFFNEY, 1984 Genetic heterozygosity and growth rate in *Mytilus edulis*. Mar. Biol. **82**: 1–7.

KOEHN, R. K., R. I. E. NEWELL and F. W. IMMERMANN, 1980 Maintenance of an aminopeptidase allele frequency cline by natural selection. Proc. Natl. Acad. Sci. USA **77**: 5385–5389.

KOEHN, R. K., W. J. DIEHL and T. M. SCOTT, 1988 The differential contribution by individual enzymes of glycolysis and protein catabolism to the relationship between heterozygosity and growth rate in the coot clam, *Mulinia lateralis*. Genetics **118**: 121–130.

MALLET, A. L., E. ZOUROS, K. E. GARTNER-KEPKAY and K. R. FREEMAN, 1986 Genetics of growth in blue mussels: family and enzyme-heterozygosity effects. Mar. Biol. **92**: 475–482.

MITTON, J. B., and M. C. GRANT, 1984 Associations among protein heterozygosity, growth rate, and developmental homeostasis. Annu. Rev. Ecol. Syst. **15**: 479–499.

NEI, M., 1987 *Molecular Evolutionary Genetics*. Columbia University Press, New York.

OHTA, T. 1971 Associative overdominance caused by linked detrimental mutations. Genet. Res. **18**: 277–286

PASTEUR, N., G. PASTEUR, F. BONHOMME, J. CATALAN and J. BRITTON-DAVIDIAN, 1988 *Practical Isozyme Genetics*. Ellis Horwood Ltd, Chichester.

PIERCE, B. A., and J. B. MITTON, 1982 Allozyme heterozygosity and growth in the tiger salamander, *Ambystoma tigrinum*. J. Hered. **73**: 250–253.

SEED, R., 1980 Shell growth and form in the bivalvia, pp. 23–67 in *Skeletal Growth of Aquatic Organisms—Biological Records of Environmental Change*, edited by D. C. RHOADS and R. A. LUTZ. Plenum Press, New York.

SMITH, S. M., 1974 Mollusca dredged off Musselburgh, Firth and Forth, Scotland in 1972, with particular reference to the population of *Spisula solida* (L.). J. Conchol. **28**: 217–224.

SMOUSE, P. E., 1986 The fitness consequences of multiple-locus heterozygosity under the multiplicative overdominance and inbreeding depression models. Evolution **40**: 946–957.

SOKAL, R. R., and F. J. ROHLF, 1981 *Biometry*. Freeman & Co., New York.

WEIR, B. S., and C. C. COCKERHAM, 1984 Estimating F-statistics for analysis of population structure. Evolution **38**: 1358–1370.

ZOUROS, E., 1993 Associative overdominance: evaluating the effects of inbreeding and linkage desequilibrium. Genetica **89**: 35–46.

ZOUROS, E., and D. W. FOLTZ, 1984 Possible explanations of heterozygote deficiency in bivalve molluscs. Malacologia **25**: 583–591.

ZOUROS, E., S. M. SINGH and H. E. MILES, 1980 Growth rate in oysters: an overdominant phenotype and its possible explanations. Evolution **34**: 856–867.

ZOUROS, E., S. M. SINGH, D. W. FOLTZ and A. L. MALLET, 1983 Post-settlement viability in the American oyster (*Crassostrea virginica*): an overdominant phenotype. Genet. Res. **41**: 259–270.

ZOUROS, E., M. ROMERO-DOREY and A. L. MALLET, 1988 Heterozygosity and growth in marine bivalves: further data and possible explanations. Evolution **42**: 1332–1341.

Communicating editor: D. CHARLESWORTH