Evolution of genetic variation for selected traits in successive breeding populations of maritime pine

Laurent Bouffier, Annie A. Raffin, Antoine Kremer

Title

Evolution of genetic variation for selected traits in successive breeding populations of maritime pine.

Authors

Laurent Bouffier [1], Annie Raffin [2], Antoine Kremer [3*]

[1] bouffier@pierroton.inra.fr

INRA, UMR1202 Biodiversity Genes & Communities, F-33612 Cestas

[2] raffin@pierroton.inra.fr

INRA, UMR1202 Biodiversity Genes & Communities, F-33612 Cestas

[3] kremer@pierroton.inra.fr

INRA, UMR1202 Biodiversity Genes & Communities, F-33612 Cestas

[*]**Corresponding author**

Antoine Kremer

antoine.kremer@pierroton.inra.fr

INRA, UMR BIOGECO, 69 Route d'Arcachon, 33612 Cestas, France

Phone : 0033.5.57.12.28.32 Fax : 0033.5.57.12.28.81

Keywords

Pinus pinaster Ait. / genetic variation / correlation / individual model / evolution / breeding population

Short title

Evolution of variation in breeding populations.

5 150 words

Abstract

Directional selection impacts a trait distribution by shifting its mean and reducing its variance. The change of variance is of major importance as the response to selection in subsequent generations is highly dependent of the genetic variability available in the population. In this contribution, evolution of genetic variation was investigated through the first breeding populations of the French maritime pine (*Pinus pinaster* Ait.) breeding program.

We considered three populations: P0 (the forest where plus trees were initially selected), G0 (the plus tree population) and G1 (the population composed of trees selected in the progenies of G0). Analyses focused on the following selected traits: total height (H), girth at 1.30 m (D) and stem deviation to verticality (S). More than 150 000 trees from 25 tests of three distinct populations were studied with an individual genetic model. Accurate genetic parameters were obtained by taking all relationships between trees into account.

For H and D, we found a strong decrease of the genetic variation from P0 to G0 corresponding to the initial selection of plus trees, which constitutes the base population of the breeding program. Then, despite the second step of selection applied, no appreciable evolution arose from comparisons between G0 and G1 for these traits. For S, the evolution is less significant as

50 phenotypic variation slightly increased, possibly due to changes of
51 silvicultural practices.
52

Introduction

The genetic variation of a population is the key factor in determining its response to natural or artificial selection and thus its evolutionary potential. In the present work, we intended to monitor the changes of the genetic variance as a result of directional selection conducted in breeding populations of maritime pine (*Pinus pinaster* Ait.), a major forest tree species growing in the southwest of France. Monitoring of genetic variation along breeding populations has been undertaken earlier , but on neutral traits in Sitka spruce (Chaisurisri and El-Kassaby, 1994) and in Douglas-fir (El-Kassaby and Ritland, 1996). In contrast to these earlier investigations, our study focuses on traits that underwent selection, hence on genetic variance and not on heterozygosity.

Many analytical analyses and simulations predict the evolution of genetic variance of a selected trait (Bulmer, 1971; Van der Werf and Boer, 1990; Verrier et al., 1991; De Rochambeau et al., 2000). Selection is expected to rapidly reduce the genetic variation which then stabilizes except in small populations where the erosion of variation continues due to genetic drift and inbreeding. However many hypotheses underlie these models and few studies have been carried out on real populations under selection (Sorensen and Hill, 1982; Meyer and Hill, 1991; Dupont-Nivet et al., 2001). Dupont-Nivet et al. (2001) observed a strong decrease of genetic variation in the two

first generations of a snail population undergoing selection followed by an equilibrium phase. Meyer and Hill (1991) reported a reduction of genetic variation during 23 generations in a population of mice selected for food intake. They concluded that the evolution of allele frequencies played a major role in that trend. Sorensen and Hill (1982) studied populations of *Drosophila* and found various patterns for the evolution of genetic variation according to the initial allele frequencies. The evolution of variability thus appears to depend both on the population considered and on the genetic basis of the trait studied.

This paper examines the evolution of genetic variation of the selected traits and their correlations in three successive populations of the French maritime pine breeding program.

Maritime pine (*Pinus pinaster* Ait.) represents one million hectares of cultivated forest in Aquitaine (southwestern France). A breeding program has been implemented since the early sixties (Durel, 1992; GIS, 2002) using a recurrent selection scheme which consists of successive cycles of selection of candidate trees and their crossings (Zobel and Talbert, 1984). The dual goals of the program have been to (1) obtain genetic gain in growth and stem straightness and (2) preserve diversity in the breeding populations. The former goal was achieved since an improvement of 15 % for volume and form was observed in the first varieties compared to unimproved material. Today genetic gains amount to 30 % in the most recent varieties

97 (GIS, 2002). The achievement of the latter goal was less studied but it is
98 considered essential because it allows for future gains and for the
99 incorporation of new selection criteria.

100 Genetic variation is estimated for the selected traits based on the
101 “individual model” (Gwaze et al., 2002), also called “animal model” as it
102 was first developed in the context of livestock breeding programs (Kennedy
103 et al., 1988). Its adaptation to trees was implemented since breeding
104 populations are moving into advanced generations (Kerr, 1998). The
105 individual model is adapted from the mixed model (Henderson, 1975). This
106 methodology takes into account all the pedigree information to accurately
107 estimate both the genetic parameters of the base population and the breeding
108 values of all genotypes by restricted maximum likelihood. Advantages of
109 the individual model compared to the more traditional least-squares analysis
110 were outlined by Lynch and Walsch (1998). First, as fixed and random
111 effects are estimated simultaneously, the precision of estimates of
112 environmental and genetic main effects is increased. Second, the individual
113 model is better suited to unbalanced data which are frequent in the case of
114 tree breeding populations, due to unpredictable mortality in long living
115 species. Third, the method takes into account phenotypic values of related
116 individuals over multiple generations and multiple progeny tests, hence
117 increasing the number of phenotypic predictors and diversifying the genetic

relatedness among trees. Finally, the individual model accounts for selection provided that all information used in selection is included in the analysis.

Durel et al. (1998) were among the first to estimate genetic parameters in a tree population with the individual model. In their study, genetic parameters were computed in an overall analysis across seven generations of apple trees. As for forest trees, the method is now widely used in the radiata pine breeding program to rank genotypes within and across generations (Jayawickrama and Carson, 2000). Others studies have also used the individual tree genetic model on more limited data sets (Gwaze et al., 2001; Gwaze et al., 2002; Dutkowski et al., 2002; Klapste et al., 2007). In the present study we apply the individual model to monitor the changes of genetic variation and correlation over three successive populations for traits that underwent repeated directional selection.

Materials and methods

Breeding populations and progeny tests

For the sake of clarity, the different populations of the maritime pine breeding program are defined as follows (Figure 1):

(i) **P0 population** is the Landes population which has proven to be the best adapted maritime pine provenance for Southwestern France (Illy, 1966). Field tests comparing different geographic seed sources, established as early as 1930, clearly showed that the local provenance exhibited the highest survival and growth potentials. Overall the whole Landes forest covers about one million hectares, with no significant population or ecotypic differentiation (Baradat and Marpeau-Bezard, 1988).

(ii) **G0 population** is the subset of 635 plus trees i.e. trees phenotypically selected during the sixties within the population P0, of which a sample of 320 were used in this study. During approximately 10 years, adult stands in the Landes forest were visited and outstanding trees in regard to stem volume and straightness were mapped and recorded by using a phenotypic index of selection. Details of the selection procedure are available in Illy (1966).

(iii) **P1 population** gathers all the progenies obtained in the subsequent improvement steps by crosses between G0 trees. The 635 G0 trees were grafted as clonal archives and subsequently crossed using various mating schemes (polycross, factorial or nested designs).

(iv) **G1 population** is the new breeding population of about 2600 trees, individually selected within P1. Index selection combining growth and straightness traits was achieved in the P1 progenies, using family and individual values as phenotypic predictors of the breeding value of selection candidates. About 5 % of the P1 trees were selected to build G1 population.

(v) **P2 population** gathers all progenies obtained by crossing G1 trees following the recurrent strategy of the breeding scheme. The 2600 trees of G1 were crossed using different mating designs (mainly polycross and nested mating designs).

Figure 1

Our analysis focuses on the following three populations: the original population (P0) and the two breeding populations obtained after a selection step (G0 and G1) because their variation can be accurately estimated with the subsequent progeny tests. We compiled data obtained from 25 progeny tests allowing to estimate the genetic variance for the selected traits: 3 progeny tests established from unselected seeds collected throughout

175 southwest France forest (to estimate the genetic variance in P0), 7 progeny
176 tests from P1 population and 15 progeny tests from P2 population. They
177 correspond to different mating designs (open pollination, factorial, nested
178 design and polycross) and each progeny test comprises on average 135
179 progenies and 9000 trees (Table 1). From here onwards, “progeny test” will
180 be called “test”. The experimental designs are either complete or incomplete
181 blocks with plot sizes varying between one to 10 trees depending on the test
182 considered. A “block” is a test subdivision comprising several “plots”, each
183 consisting of one progeny and spreads over homogeneous site conditions. A
184 block is complete when it comprises all progenies, it is incomplete when it
185 contains only a subset of progenies.

Table 1 186

187 *Measurements*

188 Two growth traits - total height (H) and girth at breast height (D) - and a
189 trait relative to stem form (S) were measured between 7 and 13 years
190 depending on the test (Table 1). Different assessments were used for S over
191 the years. Thus, we restricted our analysis of S on those tests for which the
192 same assessment was made repeatedly. The assessment consisted of
193 measuring the stem deviation to verticality, as the angle formed by the stem
194 and a virtual vertical axis passing through the base of the stem.

195

196 *Genetic model*

The individual model was used to subdivide the phenotypic value of each tree in its genetic and environmental components. As both environmental and genetic effects are computed simultaneously, the best linear unbiased predictor (BLUP) of the genotypes was obtained and the genetic parameters of the base population (i.e. highest ancestors registered) were estimated (Lynch and Walsch, 1998).

It is important to note that the genetic variation estimated with an individual model depends on the pedigree considered. When the complete multigenerational relationship matrix is considered, genetic parameters of the base population are estimated. To obtain genetic parameters of an advance population "t", only the relationship matrix computed from all individuals up to that ancestor population must be kept. The population "t" is thus assumed to be the base population (Sorensen and Kennedy, 1984; Meyer and Hill, 1991).

The following mixed-model was considered:

$$y = X.b + Z_1.a + Z_2.v + e$$

where y is the vector of observations

b is the vector of fixed effects: "test" and "block (test)"

a is the vector of genetic effect: individual additive genetic values

v is the vector of plot effect: "block(test)×progeny"

e is the vector of residuals

X, Z_1 and Z_2 are the incidence matrices linking observations to the effects

No "progeny×test" interactions were considered as only few progenies were common to different tests. Furthermore it has been shown that this effect is minor in the Landes area (Bouffier, 2007).

The random effects fit a normal distribution whose parameters were:

$$E\begin{bmatrix} a \\ v \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } Var\begin{bmatrix} a \\ v \\ e \end{bmatrix} = \begin{bmatrix} G & 0 & 0 \\ 0 & H & 0 \\ 0 & 0 & R \end{bmatrix}$$

The variance-covariance matrices were defined as follows:

$$G = A.\sigma_A^2 \qquad H = I.\sigma_v^2 \qquad R = I.\sigma_e^2$$

with: A the additive genetic relationship matrix (A was computed from a pedigree file which takes into account all the relationships between individuals)

I the identity matrix

σ_A^2 the additive genetic variance

σ_v^2 the plot variance

σ_e^2 the residual variance

The estimates of the fixed and random effects were obtained by solving Henderson's mixed model equations (Henderson, 1975) with the restricted maximum-likelihood (REML) method using the ASReml software (Gilmour et al., 2002).

As the variances are assumed to be independent, the phenotypic variance σ_P^2 is expressed as follows:

$$\sigma_P^2 = \sigma_A^2 + \sigma_v^2 + \sigma_e^2$$

Variation parameters

Univariate analyses were performed for estimating genetic and phenotypic variation. The variation of the selected traits – H, D and S – was first expressed by two widely used standardized assessments: narrow-sense heritability (h^2) and coefficient of additive genetic variation (CV_A). For comparative purposes, Houle (1991) showed that genetic variance is more appropriately standardized by the trait mean (CV_A) than by the phenotypic variance (h^2), and that heritability is rather useful for making predictions about the absolute response to selection, and CV_A for assessing genetic variation. In this study, we used both parameters and we also included the phenotypic coefficient of variation (CV_P). As our study is an overall analysis across many experimental designs established over the past 40 years, assessments of CV_P allow us to check for major environmental sources of variation that may have occurred during this period. Heritability and the two coefficients of variation were computed as follows:

$$h^2 = \frac{\sigma_A^2}{\sigma_P^2} \qquad CV_P = \frac{\sigma_P}{\bar{x}} \qquad CV_A = \frac{\sigma_A}{\bar{x}}$$

Since ASReml also provides the estimated breeding values for each parent genotype, we also computed a coefficient of variation with the breeding values (CV_{BV}) which can be considered as a third estimate of genetic variation:

$$CV_{BV} = \frac{\sigma_{BV}}{\bar{x}}$$

Correlation parameters

For correlation estimates, we considered a bivariate analysis but, because of a lack of convergence for the maximum likelihood under the full model, we decided not to include $Z_2.v$ in the model. The use of this simplified model implies that, for correlation estimates, σ_A^2 includes both additive and the plot variances. The genetic variance is thus biased upward compared to the full model.

The estimates of phenotypic (r_P) and additive genetic (r_G) correlations between pairs of traits were obtained with bivariate analyses. Genetic correlations were also estimated with the breeding values (r_{BV}) using Pearson's correlation.

Standard errors and statistical tests

279 The standard errors of h^2, σ_A^2, σ_P^2, r_P and r_G were calculated with ASReml
280 using a standard Taylor series approximation (Gilmour et al., 2002) and
281 those of CV_P , CV_A and r_{BV} were estimated with the approximation
282 proposed by Lynch and Walsh (1998).
283
284 In the following analyses, two estimates x_1 and x_2 associated with standard
285 errors σ_1 and σ_2 were considered significantly different if the statistical
286 test $u = \frac{|x_1 - x_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}}$, which is assumed to be distributed as a standard normal
287 law, was higher than 1.96 (bilateral significance level of 5 %).
288
289 *Methods used to estimate parameters of genetic variation*
290 Three methods were carried out to estimate the variation of the selected
291 traits (Table 2). All of them take advantage of the individual model
292 previously described but are based on different data sets and refer to
293 different populations.
294
295 *Method I*: each test was analysed individually considering the genetic
296 relationship matrix truncated to the parent level. We thereby obtained an
297 estimation of the genetic variation within the parental population. For
298 example, the analysis of a test from the P1 population will provide an
299 estimation of the G0 genetic parameters.

Method II: tests belonging to the same population were analysed all together (and not individually as for Method I) considering also the genetic relationship matrix truncated to the parent level. Thus, estimates of genetic variation within P0 (respectively G0, G1) were obtained from data of tests 11 to 13 (respectively 22 to 28, 31 to 47). Similarly, breeding values of P0, G0 and G1 trees and their coefficient of variation (CV_{BV}) can be estimated from their progenies.

Method III: a global analysis was performed with all the tests except tests 11 to 13 because no pedigree connection existed between the population P0 and the following ones. The pedigree considered took into account the complete multigeneration genetic relationships. We thus obtained another estimate for the G0 genetic parameters (considered, in this analysis, as the base population) and for CV_{BV} of populations G0 and G1.

The same three methods were used to estimate correlations between the selected traits (Table 2).

Table 2

Results

1. Analyses per test (Method I)

Figure 2 displays the results of each single test analysis for H, D and S. There is a wide range of variation of each parameter (h^2, CV_A and CV_P) among tests within a population. For example, CV_A of D (Figure 2 b2) varies between 5.1 to 14.1 % (mean standard errors is 0.9 %) among the different tests of population G1. For a given test and trait, CV_{BV} is always lower than CV_A.

In most of the tests, H appears to be slightly more heritable and exhibits a lower genetic coefficient of variation than D. The phenotypic coefficient of variation is clearly lower for H (between 8.4 and 18.3 %) than D (between 14.5 and 34.7 %). Heritability of S is about of the same magnitude than H or D but the genetic and phenotypic coefficients of variation are much higher (CV_A superior to 20 % and CV_P superior to 50 %).

Variation parameters of G0 and G1 are more accurate (ie. lower standard errors) than those of P0 as they are estimated with tests implying more progenies.

Figure 2

Correlations were estimated for each single test (Figure 3). Note that S is the deviation to verticality, thus a positive correlation between S and a growth trait (H or D) means that straightness is unfavourably correlated with

341 growth. Phenotypic correlations are high between H and D and moderate
342 between S and growth traits. Genetic correlations are higher than the
343 phenotypic ones albeit estimated with a larger standard error. As for the
344 genetic variance, there is a wide range of variation of correlations among the
345 different tests.

Figure 3 346

347 *2. Analyses per population (Method II)*

348 Variation parameters estimated for each entire population with the genetic
349 relationship matrix truncated to the parent level are presented in Figure 2
350 (values linked by a line) and Table 3. For each population, a large set of data
351 was analysed: 6 105 trees were considered for P0, 67 223 trees for G0 and
352 86 582 trees for G1. As a consequence, variation parameters were estimated
353 with low standard errors compared to the estimates from Method I.
354 The evolution of genetic variation over breeding populations is similar for H
355 and D. CV_A decreases between P0 and G0 (from 10.2 % to 5.6 % for H, and
356 from 11.1 % to 7.4 % for D) then remains constant between G0 and G1.
357 Likewise CV_{BV} of these two traits decreases substantially from P0 to G0
358 (from 8.2 % to 5.3 % for H, and from 8.0 % to 5.5 % for D) then slightly
359 from G0 to G1 (from 5.3 % to 4.1 % for H, and from 5.5 % to 4.1 % for D).
360 While CV_P is rather stable from P0 to G0, it tends to decrease between G0
361 and G1. Heritability, which follows the same pattern than CV_A from P0 to
362 G0, slightly increases from G0 to G1.

363 The evolution of genetic parameters for S is similar between P0 and G0 but
364 the trend is weaker and not significant: CV_A decreases from 24.1 % to
365 20.9 % and CV_{BV} from 18.0 % to 17.0 %. Then an increase is observed
366 between G0 and G1 (from 20.9 % to 26.9 % for CV_A and from 17.0 % to
367 20.2 % for CV_{BV}). CV_P slightly increases from P0 to G1.

Table 3 368

369 Method II was also used to estimate the correlations between the three
370 selected traits (Figure 3 and Table 4). No strong pattern can be observed in
371 regard to the standard errors. Nevertheless the genetic correlation between H
372 and D is slightly lower in breeding populations G0 and G1 than in P0. There
373 is no significant change of the correlation between H and S, and between D
374 and S (Table 4).

Table 4 375

376

377 *3. Analysis of the whole data (Method III)*

378 Method III provides an overall estimation of the genetic variation in G0
379 across all populations and tests, by taking into account multigeneration
380 genetic relationships. However this method does not allow the estimation of
381 parameters of P0, as trees from tests 11 to 13 are not genetically related to
382 trees of subsequent populations. Method III is based on a very large sample
383 of trees: for example, 153 805 trees were considered to estimate heritability
384 of H. Stem deviation from verticality (S) exhibits greater genetic variation

(both in terms of heritability and genetic variance) than H or D (Table 3). Estimates of variation of G0 are similar between Method III (analysis based on P1 and P2 populations) and Method II (analysis based only on P1 population). However, for H and D, the heritability is slightly higher and the phenotypic variation lower when both populations P1 and P2 are considered.

Method III also facilitates the estimation of CV_{BV} in G0 and G1 with the same data set. No strong evolution from G0 to G1 is highlighted: CV_{BV} of H slightly decreases, no significant change is found for D and CV_{BV} of S slightly increases. Thus these results confirm those obtained by Method II.

Correlation estimates among traits in G0 population are very similar between Method II and Method III (Table 4). Correlations were also estimated within G0 and G1 trees with the breeding values. Only a slight decrease was observed from G0 to G1 for H–D correlation.

Discussion

1. *Level of genetic variation for growth traits (H, D) and stem deviation to verticality (S)*

Our study shows that genetic variation of these traits in natural populations and in the very early breeding populations is moderate (Table 3). Cornelius (1993) compiled genetic parameter (h^2 and CV_A) estimates from 67 published papers, mainly on *Pinus* species. They were based on experimental designs established with progenies of selected trees from natural populations (corresponding thus to our tests with P1 trees) and can therefore be compared to our results. Most of the heritability estimates (respectively CV_A) of H and D varied between 0.05 and 0.40 (respectively between 5 and 15 %). Results from our individual tests are within this range but our study clearly indicates that estimates can be quite variable across tests, suggesting that they are highly dependent on the sampling of genetic entries and the site conditions (Figure 2). Consequently, many authors have tried to compile data from several tests to estimate the genetic parameters of a population more accurately. The heritability of growth traits of the base population of *Pinus elliottii* breeding program (G0 population) varied between 0.12 and 0.16 over a large set of tests (Dieters et al., 1995; Hodge and White, 1992). Jayawickrama (2001) analysed more than 150 000 radiata pines to estimate genetic variation of the plus tree population (G0

424 population): heritability amounted to 0.11 for girth, 0.13 for height and 0.19
425 for straightness. These three analyses based on large data sets reported low
426 to moderate genotype × environment interactions. Our estimates for
427 population G0 and G1 are of similar magnitude (Table 3 - Method II). As
428 expected, the unselected population P0 displays higher estimates than
429 populations under selection. The high variation across tests (Figure 2)
430 further indicates a need for multiple tests to reliably estimate genetic
431 variation. In this respect, the individual model is a recommended method ,
432 as suggested by the decrease of the standard error of variation parameters as
433 we moved from Method I to Method III (data not shown).
434
435 Regarding the comparison of the level of genetic variation among
436 traits, Cornelius (1993) concluded that H is more heritable than D (0.25 for
437 the median heritability of height vs. 0.19 for diameter), and exhibits higher
438 genetic variance (as shown by the coefficient of genetic variation). While
439 we draw similar conclusions for heritability (h^2 higher for H than for D),
440 there is an opposite trend for the coefficients of variation: both CV_A and
441 CV_P are higher for D than H (Table 3). Lending support to our result,
442 Gwaze et al. (2001) also observed higher heritability and a lower coefficient
443 of variation for height compared to diameter in *Pinus taeda* based on an
444 individual tree model. The coefficients of variation of H and D can be

compared in our study without corrections because they have the same dimensionality (Houle, 1991).

Genetic parameters of S are more difficult to compare across studies because several different phenotypic assessments were used to assess stem straightness. However, as for other studies (Cornelius, 1993; Jayawickrama, 2001), a higher genetic variation is observed compared to the growth traits (Figure 2, Table 3).

2. Evolution of the genetic variation of selected traits throughout successive populations

Regardless of the method used, phenotypic variation shows no clear evolution for H and D from P0 to G0 but decreases slightly from G0 to G1 (Figure 2, Table 3). On the contrary, there is a clear decrease of genetic variation from P0 to G0 and a very minor decrease from G0 to G1. The pattern of genetic variation is consistent among the parameters used (CV_A or CV_{BV}). We will therefore restrict the discussion to CV_{BV} as CV_{BV} was the only parameter that could be used for comparing genetic variation over breeding populations with Method III. CV_{BV} is reduced on average by 35 % for H and by 31 % for D between P0 and G0 (Table 3).

There is still a significant decrease of CV_{BV} of growth traits from G0 and G1 according to Method II (by 23 % for H and by 25 % for D, see Table 3).

Yet, according to Method III, the most accurate analysis as it gathers both P1 and P2 populations, the decrease is weak for H (12 %) and not significant for D between populations G0 and G1 (Table 3).

In conclusion, while mean values of the two growth traits increased as a result of directional selection from G0 to G1, only a very slight reduction of genetic variation was observed. Similarly, King et al. (1998) reported no change of genetic variation between two breeding populations of *Pinus radiata* equivalent to the ones we referred to as G0 and G1.

Genetic variance of S follows a similar trend across generations albeit less pronounced: there is a non-significant decrease from P0 to G0 followed by a slight increase from G0 to G1 (Table 3). The increase from G0 to G1 is unexpected as stem straightness underwent recurrent directional selection like the two other traits. However, silvicultural pratices have changed over time, and may have impacted stem straightness more than the other two growth traits. Intensive treatments such as the use of fertilizers and ploughing that became more frequent in recent times, may have increased the environmental and genetic variation of traits. This is suggested by the larger increase of the phenotypic variance of S from P0 to G1 (Figure 2, Table 3), in contrast to H and D. Under such circumstances, one may suspect that the genetic variance has been impacted as well, blurring the effect of directional artificial selection that we tried to monitor.

Changes of genetic variance in artificial breeding populations may result from either drift effects due to the reduction of population size, or from directional selection. As the pedigree is known over two generations, the "status number" (N_S) (Lindgren et al., 1996) can be used to provide an estimate of the population effective size. N_s is "the number of unrelated and non-inbred genotypes in an ideal panmictic population, which is expected to produce offspring with the same coefficient of inbreeding as the progeny of the considered population following random mating" (Lindgren et al., 1997). Based on the pedigree data of the tests considered in this study, the status number of G1 amounts to about 90 which can be compared to the 320 unrelated plus trees of the G0 population analysed here. The estimated decrease of genetic variance at generation "t" due to the reduction of population size should amount to $\left(1-\frac{1}{2\,Ne}\right)$ of the genetic variance at generation "t-1" where Ne is the effective size (Lynch and Walsh, 1998). Consequently, the reduction of genetic variance due to the reduction of population size remains extremely small, and is therefore most likely to be caused by directional selection.

The evolution of genetic variation in populations undergoing selection was first investigated by Bulmer (1971). By considering a quantitative trait controlled by an infinite number of loci, he subdivided genetic variance (V_A) into two components: the "equilibrium genetic variance" also called

genic variance (the first term of the following equation) and the "disequilibrium contribution" (the second term):

$$V_A = \sum_i Var(g_i) + \sum_{i \neq j} Cov(g_i, g_j)$$

with $Var(g_i)$ the variance at the i^{th} locus and $Cov(g_i, g_j)$ the covariance between the i^{th} and the j^{th} loci.

Under this model, Bulmer (1971) showed that directional selection induces a negative disequilibrium contribution, and thus the genetic variance decreases over generations. This effect, known as the "Bulmer effect", is temporary and the disequilibrium contribution progressively approaches zero if selection is relaxed. Bulmer (1971) showed with an analytical model that the decrease of genetic variance under selection is high in the first generations and rapidly stabilised. The equilibrium stage occurs when the effects of selection and recombination counterbalance each other. However, the reduction of the genetic variance can be inflated by the reduction of genic variance. Indeed, the genic variance decreases if the trait under selection is determined by a finite number of genes or if small populations are considered (De Rochambeau et al., 2000).

The reduction of genetic variation observed between P0 and G0 is thus mainly explained by the phenotypic selection conducted in the 1960's to constitute the "plus" trees population, and may be due to the Bulmer effect. A mass selection was performed throughout the Landes forest, using a

procedure that permitted the consideration of a genetic component in the phenotypic superiority of selected trees. The method was based on the standardized value of a candidate tree compared to its 30 immediate neighbours (Illy, 1966), thus taking into account environmental effects. Despite the moderate heritability of the selected traits (from 0.20 to 0.50, see Table 3), the extremely high selection rate that was used during the mass selection was sufficient to reduce the genetic variation in the subsequent generation (G0). Illy (1966) reported that one tree out of 70 000 was selected during this selection step (this estimation is based on the number of trees screened for plus trees selection).

As a genetic selection step was then achieved to build the G1 population from progenies of G0 trees, we also expected a significant decrease of genetic variation for selected traits. However the decrease was much lower or non significant depending on the method considered. Various hypotheses can be suggested to interpret theses results. First, the accuracy of the analyses may not be able to detect a slight decrease of genetic variation. Second, the selection was performed on three criteria (H, D and S); therefore the selection intensity for each of them may be more limited and may have been much lower than the selection intensity used during the first stage (from P0 to G0). Third, the equilibrium phase may be achieved after the selection of the plus trees but this hypothesis is unlikely as only one selection step was performed to obtain the population G0.

553

554 *3. Level and evolution of the correlations between the selected*
555 *traits throughout successive populations*

556 Genetic correlations are highly positive (favourable) between H and D while
557 they are slightly positive (unfavourable) between growth traits and stem
558 deviation to verticality (Figure 3, Table 4). Considering both Method II and
559 Method III, no consistent pattern in the correlation change was found,
560 except for a slight decrease between P0 and G0 for H - D correlation.
561 The evolution of genetic correlation was investigated by simulations
562 according to the relative weights of index selection and to the initial
563 variation of the traits (McMillan et al., 1995). If the initial genetic
564 correlation is positive, simulations suggest a decline towards zero, the rate
565 of change increasing with the heritability of one or both traits. If genetic
566 correlation is negative, there are two contrasting trends: either the
567 correlation increases towards zero if economic weights are unequal or it
568 declines to –1 if they are equal. The slight change of the H - D correlation
569 observed is in agreement with these simulations. As economic weights of
570 growth traits and straightness are similar, we expect an increase of the D - S
571 correlation through the breeding populations. Indeed we found a slight
572 increase for r_{BV} but it appears non significant.

573

574 *4. Conclusion*

575 Our results showed that even if the population effective size has been
576 substantially reduced over successive breeding populations, the genetic
577 variance for the selected traits has not followed the same trend. Indeed, after
578 a decrease when selecting for plus trees, the genetic variation remained
579 fairly constant suggesting the possibility to maintain genetic gains over
580 future generations with this recurrent selection strategy.
581 Lastly we may suppose that genetic variation of unselected traits has been
582 maintained above the level we observed for selected traits. Selection for
583 new criteria could therefore be implemented at the level of G2 without
584 enriching the genetic variation from external genetic resources, provided
585 that genetic correlation between the new criteria and growth or straightness
586 remains low.
587

588 **Acknowledgements**

589

590 We thank Florence Jaffrezic from INRA Jouy-en-Josas for help using the
591 individual model. Plantations were maintained by the members of the GIS
592 “Pin Maritime du Futur” (FCBA, CRPF, CPFA, ONF, INRA). We thank
593 also the Experimental Unit of INRA for field measurements.
594 This work was supported by funding from the French Ministry of
595 Agriculture and the Région Aquitaine.

596

597 **References**

598 Baradat P, Marpeau-Bezard A (1988). Le pin maritime (*Pinus pinaster*
599 Ait.): biologie et génétique des terpènes pour la connaissance et
600 l'amélioration de l'espèce. PhD Thesis, Université Bordeaux I.

601 Bouffier L (2007). Évolution de la variabilité génétique dans les populations
602 d'amélioration du pin maritime (*Pinus pinaster* Ait.) et conséquences
603 pour la sélection. PhD Thesis, Université Bordeaux.

604 Bulmer MG (1971). The effect of selection on genetic variability. *Am Nat*
605 **105**: 201-211.

606 Chaisurisri K, El-Kassaby YA (1994). Genetic diversity in a seed
607 production population vs. natural populations of Sitka Spruce. *Biodiv*
608 *Conserv* **3**: 512-523.

609 Cornelius J (1993). Heritabilities and additive genetic coefficients of
610 variation in forest trees. *Can J For Res* **24**: 372-379.

611 De Rochambeau H, Fournet-Hanocq, Vu Tien Khang J (2000). Measuring
612 and managing genetic variability in small populations. *Ann Zootech*
613 **49**: 77-93.

614 Dieters MJ, White TL, Hodge GR (1995). Genetic parameter estimates for
615 volume from full-sib tests of slash pine (*Pinus elliottii*). *Can J For Res*
616 **25**: 1397-1408.

617 Dupont-Nivet M, Mallard J, Bonnet JC, Blanc JM (2001). Evolution of
618 genetic variability in a population of the edible snail, *Helix aspersa*

619 Müller, undergoing domestication and short-term selection. *Heredity*
620 **87**(2): 129-135.
621 Durel CE (1992). Gain génétiques attendus après sélection sur index en
622 seconde génération d'amélioration du pin maritime. *Rev For Fr*
623 **44**(4):341-355.
624 Durel CE, Laurens F, Fouillet A, Lespinasse Y (1998). Utilization of
625 pedigree information to estimate genetic parameters from large
626 unbalanced data sets in apple. *Theor Appl Genet* **96**:1077-1085.
627 Dutkowski GW, Costa e Silva J, Gilmour AR, Lopez GA (2002). Spatial
628 analysis methods for forest genetic trials. *Can J For Res* **32**: 2201-
629 2214.
630 El-Kassaby YA, Ritland K (1996). Impact of selection and breeding on the
631 genetic diversity in Douglas-fir. *Biodiv Conserv* **5**: 795-813.
632 Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R (2002).
633 ASReml User Guide Release 1.0 VSN International Ltd, Hemel
634 Hempstead, HP1 1ES, UK.
635 GIS, collective work, Alazard P, Canteloup D, Cremiere L, Daubet A,
636 Lesgourgues Y, Merzeau D, Pastuszka P, Raffin A (2002). Genetic
637 breeding of the maritime pine in Aquitaine. GIS Work Report.
638 Gwaze DP, Bridgwater FE, Byram TD, Lowe WJ (2001). Genetic parameter
639 estimates for growth and wood density in Loblolly pine (*Pinus taeda*
640 L.). *Forest Genetics* **8**(1): 47-55.

641 Gwaze DP, Harding KJ, Purnell RC, Bridgewater FE (2002). Optimum
642 selection age for wood density in loblolly pine. *Can J For Res* **32**:
643 1393-1399.
644 Henderson CR (1975). Best linear unbiaised estimation and prediction under
645 a selection model. *Biometrics* **31**: 423-447.
646 Hodge GR, White TL (1992). Genetic parameter estimates for growth traits
647 at different ages in slash pine and some implications for breeding.
648 *Silvae Genet* **41**(4-5): 252-262.
649 Houle D (1991). Comparing evolvability and variability of quantitative
650 traits. *Genetics* **130**: 195-204.
651 Illy G (1966). Recherches sur l'amélioration génétique du pin maritime. *Ann*
652 *Sci For* **23**: 757-948.
653 Jayawickrama KJS (2001). Genetic parameter estimates for radiata pine in
654 New Zealand and New South Wales: a synthesis of results. *Silvae*
655 *Genet* **50**(2): 45-53.
656 Jayawickrama KJS, Carson MJ (2000). A breeding strategy for the New
657 Zealand radiata pine breeding cooperative. *Silvae Genet* **49**(2): 82-90.
658 Kennedy BW, Schaeffer LR, Sorensen DA (1988). Genetic properties of
659 animal models. *J Dairy Sci* **71**(2): 17-26.
660 Kerr RJ (1998). Asymptotic rates of response from forest tree breeding
661 strategies using best linear unbiased prediction. *Theor Appl Genet* **96**:
662 484-493.

King JN, Carson MJ, Johnson GR (1998). Analysis of disconnected diallel mating designs: II – Results from a third generation progeny test of the New Zealand radiata pine improvement programme. *Silvae Genet* **47**(2-3): 80-87.

Klapste J, Lstiburek M, Kobliha J (2007). Initial evaluation of half-sib progenies of Norway spruce using the best linear unbiased prediction. *J For Sci* **53**(2): 41-46.

Lindgren D, Gea L, Jefferson P (1996). Loss of genetic diversity monitored by status number. *Silvae Genet* **45**(1): 52-59.

Lindgren D, Gea LD, Jefferson PA (1997). Status number for measuring genetic diversity. *Forest Genetics* **4**(2): 69-76.

Lynch M, Walsh B (1998). *Genetics and analysis of quantitative traits*, Sinauer Associates: Sunderland, Massachussetts.

Meyer K, Hill WG (1991). Mixed model analysis of a selection experiment for food intake mice. *Genet Res* **57**: 71-81.

McMillan I, Fairfull RW, Friars GW, Quinton M (1995). The effect of simultaneous selection on the genetic correlation. *Theor Appl Genet* **91**: 776-779.

Sorensen DA, Hill WG (1982). Effect of short term directional selection on genetic variability: experiments with *Drosophila melanogaster*. *Heredity* **48**(1): 27-33.

684 Sorensen DA, Kennedy BW (1984). Estimation of genetic variances from
685 unselected and selected populations. *J Anim Sci* **59**(5): 1213-1223.
686 Van der Werf JHJ, Boer IJM (1990). Estimation of additive genetic variance
687 when base populations are selected. *J Anim Sci* **68**: 3124-3132.
688 Verrier E, Colleau JJ, Foulley JL (1991). Methods for predicting response to
689 selection in small populations under additive genetic models: a review.
690 *Livest Prod Sci* **29**: 93-114.
691 Zobel B, Talbert J (1984). *Applied forest tree improvement.* Wiley, New
692 York.
693

Figure 1

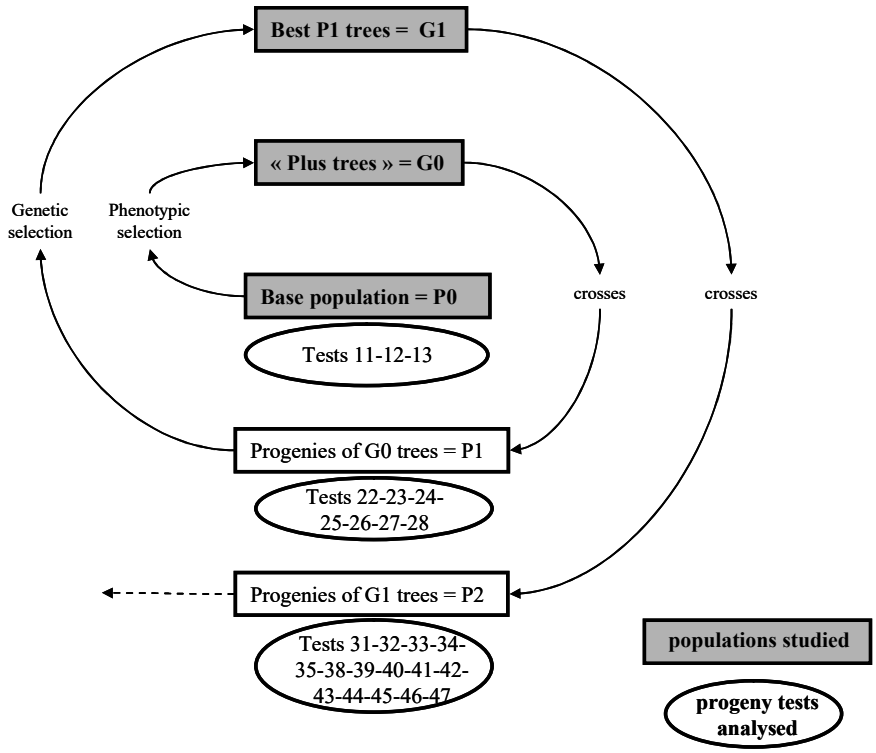
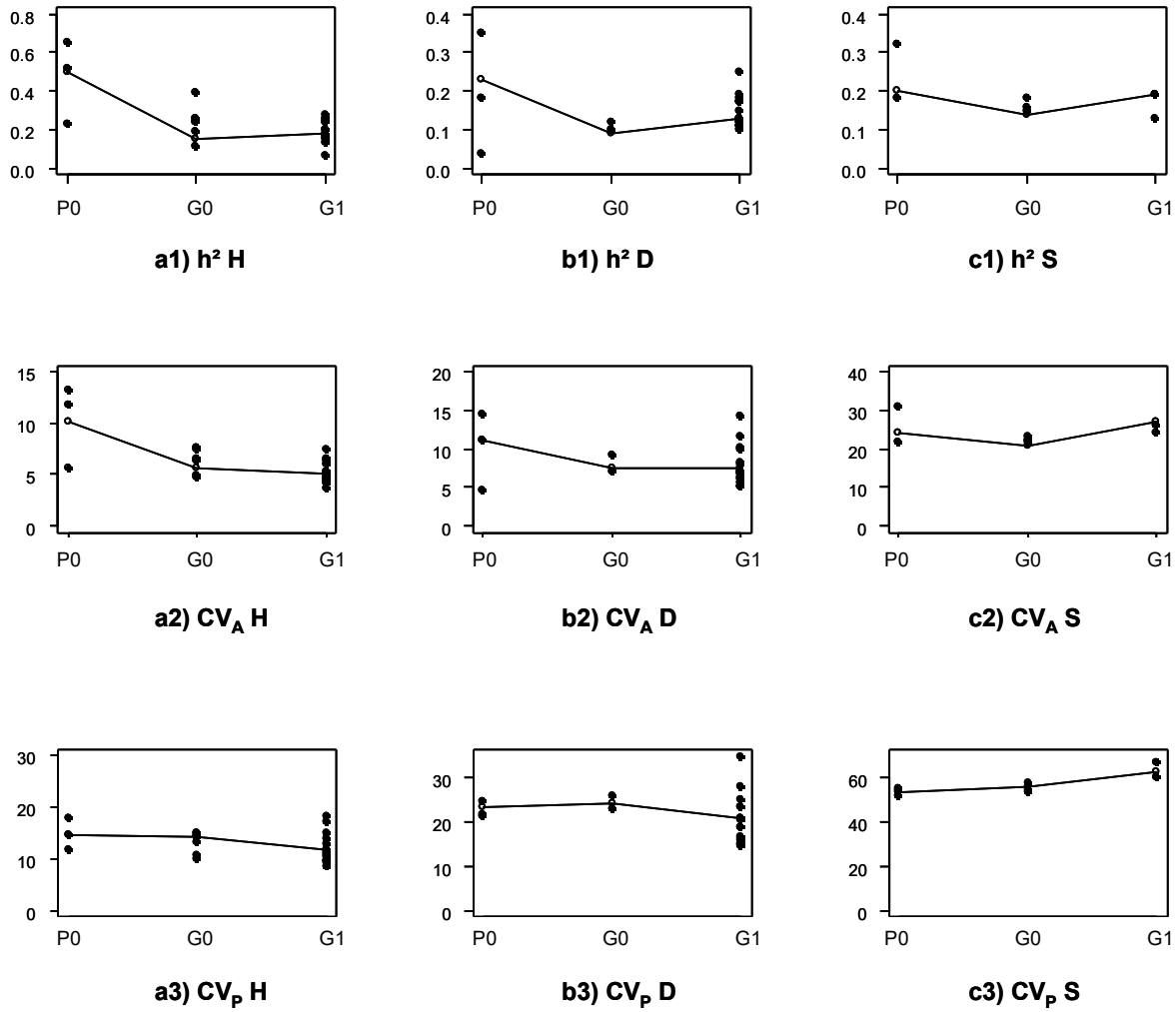
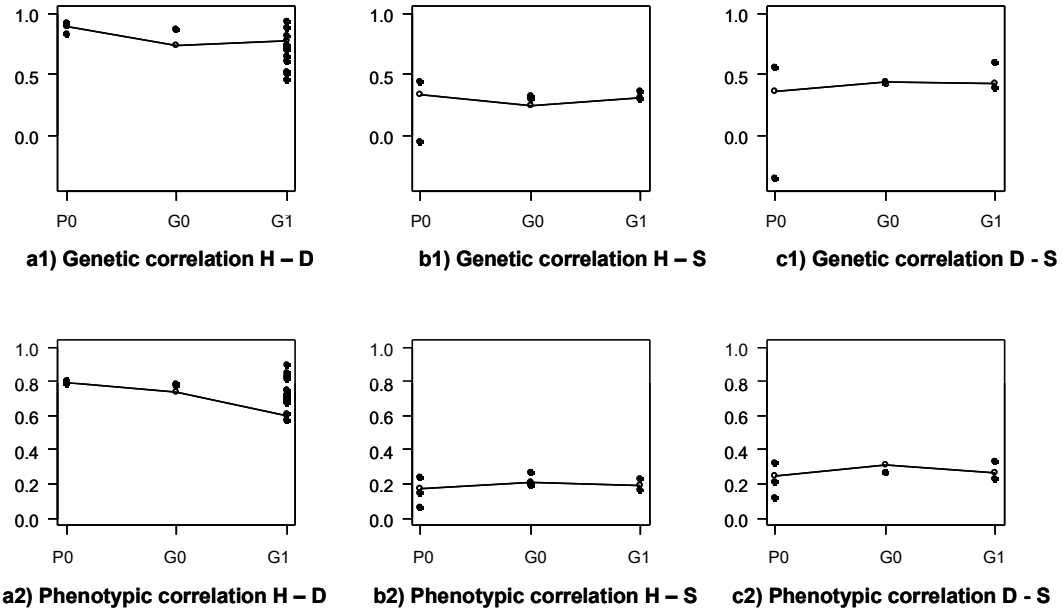

Best P1 trees = G1
« Plus trees » = G0
Genetic selection
Phenotypic selection
Base population = P0
Tests 11-12-13
crosses
crosses
Progenies of G0 trees = P1
Tests 22-23-24-25-26-27-28
Progenies of G1 trees = P2
Tests 31-32-33-34-35-38-39-40-41-42-43-44-45-46-47
populations studied
progeny tests analysed

Figure 2



(H = total height ; D = girth at breast height ; S = stem deviation to verticality ; h² = heritability ; CV_A = additive coefficient of variation ; CV_P = phenotypic coefficient of variation)

Figure 3



(H = total height ; D = girth at breast height ; S = stem deviation to verticality)

744 **Table 1**

745 Test features.

746

Population	Code	Design	Progenies	Trees	Block	Trees per plot	Plantation year	Measurement age	Height mean	Girth mean	Deviation to verticality mean
P0	11	open pollination	53	1430	5	3 to 20	1972	8	437.5	29.2	13.2
	12	open pollination	50	1288	5	8 to 12	1973	10	661.9	33.1	10.0
	13	open pollination	72	3388	5	3 to 19	1974	10	721.0	33.6	11.0
G0	22	open pollination	56	15040	90	10	1965	7	632.5	-	-
	23	factorial	144	16420	48	4 to 12	1968	7	490.1	-	-
	24	factorial	169	49769	8	9	1969	9	678.8	-	-
	25	polycross	261	21663	50	3 to 9	1975	8	583.4	31.0	-
	26	nested design	236	11745	110	4	1976	9	654.5	30.7	7.7
	27	nested design	75	3465	33	4	1977	10	707.0	-	8.0
	28	nested design	76	4132	76	8	1978	10	711.0	-	9.4
G1	31	nested design	72	5598	68	1	1981	8	572.0	26.1	4.3
	32	nested design	28	2777	97	1	1982	12	737.8	30.6	-
	33	nested design	157	14728	50	1	1985	12	920.6	43.3	-
	34	nested design	66	3966	72	1	1986	13	1023.0	47.7	-
	35	polycross	213	9188	125	4	1982	9	764.7	33.2	6.3
	38	polycross	129	6046	5	10	1992	11	937.4	43.9	-
	39	polycross	101	3535	35	1	1994	8	644.0	37.1	-
	40	polycross	101	3535	35	1	1995	8	658.3	35.6	-
	41	polycross	101	3299	35	1	1995	8	679.5	36.3	-
	42	polycross	211	7420	35	1	1995	8	766.8	40.1	-
	43	polycross	211	7455	35	1	1995	8	633.0	33.2	-
	44	polycross	211	7355	35	1	1995	8	695.7	36.4	-
	45	polycross	197	6895	35	1	1996	8	539.8	30.5	-
	46	polycross	197	5495	35	1	1996	8	644.7	32.9	-
	47	polycross	197	6160	35	1	1996	8	731.5	41.3	-

747

748

749

750

751 **Table 2**

752 Methods to estimate variability and correlations in multigenerational

753 populations using the individual model.

754

	ANALYSIS		VARIABILITY			CORRELATIONS		
	Data analysed	**Pedigree considered**	**h^2 and CV_A estimated**	**CV_{BV} estimated**	**Results**	**r_P and r_G estimated**	**r_{BV} estimated**	**Results**
Method I	test 11	parent level	P0	P0	Figure 2	P0	P0	Figure 3
	…	…	…	…		…	…	
	test 22	parent level	G0	G0		G0	G0	
	…	…	…	…		…	…	
	test 31	parent level	G1	G1		G1	G1	
	…	…	…	…		…	…	
Method II	P0 (tests 11 to 13)	parent level	P0	P0	Figure 2 and Table 3	P0	P0	Figure 3 and Table 4
	P1 (tests 22 to 28)	parent level	G0	G0		G0	G0	
	P2 (tests 31 to 47)	parent level	G1	G1		G1	G1	
Method III	P1 and P2 (tests 22 to 47)	all relationships between individuals	G0	G0 and G1	Table 3	G0	G0 and G1	Table 4

755

756 (CV_A = additive coefficient of variation ; CV_{BV} = coefficient of variation of the breeding values ; r_P =

757 phenotypic correlation ; r_G = genetic correlation ; r_{BV} = correlations estimated with breeding values)

758

759

Table 3

Variability estimated per population considering either the pedigree relationships up to the parent level (Method II) or the all pedigree relationships. (standard errors given in brackets).

For a given trait and a given parameter, different letters indicate significant difference between estimates.

		Method II						Method III			
		P0		G0		G1		G0		G1	
H	h^2	0.50 (0.08)	*a*	0.15 (0.01)	*b*	0.18 (0.01)	*c*	0.19 (0.01)		-	
	CV_A	10.2 (0.9)	*a*	5.6 (0.3)	*b*	5.0 (0.2)	*b*	5.5 (0.2)		-	
	CV_{BV}	**8.2** (0.4)	*a*	**5.3** (0.2)	*b*	**4.1** (0.1)	*c*	**5.0** (0.2)	*a'*	**4.4** (0.1)	*b'*
	CV_P	14.5 (0.1)	*a*	14.1 (0.1)	*b*	11.9 (0.0)	*c*	12.6 (0.0)		-	
D	h^2	0.23 (0.05)	*a*	0.09 (0.01)	*b*	0.13 (0.01)	*c*	0.14 (0.01)		-	
	CV_A	11.1 (1.1)	*a*	7.4 (0.4)	*b*	7.5 (0.2)	*b*	8.1 (0.2)		-	
	CV_{BV}	**8.0** (0.4)	*a*	**5.5** (0.2)	*b*	**4.1** (0.1)	*c*	**5.8** (0.2)	*a'*	**6.2** (0.1)	*a'*
	CV_P	23.2 (0.2)	*a*	24.2 (0.1)	*b*	20.9 (0.1)	*c*	21.6 (0.1)		-	
S	h^2	0.20 (0.04)	*a*	0.14 (0.02)	*a*	0.19 (0.02)	*a*	0.16 (0.01)		-	
	CV_A	24.1 (2.3)	*a*	20.9 (1.3)	*a*	26.9 (1.7)	*b*	23.2 (1.1)		-	
	CV_{BV}	**18.0** (1.0)	*a*	**17.0** (0.7)	*a*	**20.2** (0.8)	*b*	**17.2** (0.7)	*a'*	**21.2** (0.8)	*b'*
	CV_P	53.8 (0.5)	*a*	55.9 (0.4)	*b*	62.5 (0.5)	*c*	58.7 (0.3)		-	

(H = total height ; D = girth at breast height ; S = stem deviation to verticality ; h^2 = heritability ; CV_A = additive coefficient of variation ; CV_P = phenotypic coefficient of variation ; CV_{BV} = coefficient of variation of the breeding values)

Table 4

Correlations between selected traits in the successive populations (standard errors given in brackets).

For a given parameter, different letters indicate significant difference between estimates.

		Method II			Method III	
		P0	G0	G1	G0	G1
Correlations H - D	r_P	0.79 (0.01) *a*	0.74 (0.00) *b*	0.60 (0.00) *c*	0.73 (0.00)	-
	r_G	0.89 (0.01) *a*	0.73 (0.03) *b*	0.77 (0.03) *b*	0.67 (0.02)	-
	r_{BV}	0.91 (0.03) *a*	0.81 (0.03) *b*	0.85 (0.02) *b*	0.78 (0.04) *a'*	0.68 (0.02) *b'*
Correlations H - S	r_P	0.17 (0.01) *a*	0.21 (0.01) *b*	0.19 (0.01) *b*	0.20 (0.01)	-
	r_G	0.33 (0.11) *a*	0.24 (0.07) *a*	0.31 (0.07) *a*	0.27 (0.05)	-
	r_{BV}	0.37 (0.07) *a*	0.23 (0.06) *a*	0.27 (0.05) *a*	0.23 (0.06) *a'*	0.32 (0.05) *a'*
Correlations D - S	r_P	0.25 (0.01) *a*	0.31 (0.01) *b*	0.26 (0.01) *a*	0.29 (0.01)	-
	r_G	0.36 (0.12) *a*	0.44 (0.08) *a*	0.43 (0.07) *a*	0.42 (0.05)	-
	r_{BV}	0.37 (0.07) *a*	0.27 (0.05) *a*	0.38 (0.05) *a*	0.34 (0.05) *a'*	0.47 (0.05) *a'*

(H = total height ; D = girth at breast height ; S = stem deviation to verticality ; r_P = phenotypic correlation ; r_G = genetic correlation ; r_{BV} = correlations estimated with breeding values)