Detecting differential viability selection between environments by analysis of compositional differentiation at different levels of genetic integration

Elizabeth M. Gillet, [1*] Martin Ziehe [1] and Hans-Rolf Gregorius [1,2]

[1] Abteilung Forstgenetik und Forstpflanzenzüchtung, Fakultät für Forstwissenschaften und Waldökologie, Georg-August-Universität Göttingen, Büsgenweg 2, 37077 Göttingen, Germany
[2] Institut für ökologische und Populationsgenetik, Am Pfingstanger 58, 37075 Göttingen, Germany

* **Corresponding author:** egillet@gwdg.de

Abstract

Viability selection can be detected directly in an environment when the genotypes of the individuals at one ontogenetic stage (e.g. seeds) and the genotypes of the survivors at a later stage are both known, but genotypes at the earlier stage often cannot be determined. In this case, differential viability selection between environments can be detected as differences in the distributions of genetic types among survivors growing in different environments, provided that the survivors stem from random samples of seeds from the same base population (e.g. seed lot). Since common F_{ST}-outlier methods for detecting selected gene loci use only allele frequencies, selection that affects the higher hierarchical levels of genetic integration (single- or multi-locus genotypes) without changing allele frequencies is not noticed. A new method for detecting differential viability selection at any level of genetic integration enables discovery of elementary mechanisms of selection that older methods miss. It is based on two measures of compositional differentiation between environments. δ_{SD} measures qualitative differences between distributions of genetic types at any given integration level without regarding differences in their constituent alleles, while Δ_{SD} measures quantitative differences between the same distributions by additionally considering the genic differences. The difference between these measures expresses the degree to which the patterns of gene association in the genotypes differ between environments. The P-values of all measures are estimated by permutation analysis under the assumption that survivors were randomly assigned to environments. Significance indicates the occurrence of differential viability selection at the loci. As a case study, a field study of viability in juvenile beech (*Fagus sylvatica* L.) for twelve enzyme loci is reanalyzed. It turns out that the significant differential selection for genotypes detected at three loci can be attributed to three combinations of selective effects: on alleles only (SKDH-A), mostly alleles but also association patterns (LAP-A); interaction of effects on alleles and association patterns that are non-significant when viewed separately (AAT-B).

Keywords: *Compositional differentiation, gene association, genetic distance, genetic integration, genetic variation, viability selection, European beech (Fagus sylvatica L.)*

Introduction

In forestry, it is common practice to harvest seeds in large quantities in one location ("geprüfte Saatgutbestände", seed orchards) and sow them in other locations for purposes of reforestation. Mortality is typically high in the early stages of growth. Whereas a random selection of seedlings presumably succumb to non-genetic factors such as grazing or landing in generally un-inhabitable micro-environmental conditions, the mortality of some of the seedlings is expected to be due to genetic maladaptation to the environment. The survivors, by the simple fact that they survived, have proven themselves to be sufficiently genetically adapted to their environment, at least up to the time of observation. Foresters take advantage of this form of natural selection by sowing more seeds as compensation for mortality, in the expectation that the survivors will form a well-adapted, vital stand.

A gene locus is adaptive in an environment if the survival rate, or *viability*, of individuals differs between the genotypes at this locus. As a result, the distribution (i.e., relative frequencies) of the genotypes at this locus among the survivors differs from that among the initially sown seeds. Such a locus may itself be functionally involved in the survival process. It may,

however, be stochastically associated with a functionally involved locus (associated selection), in which case the locus may be considered as an indicator of selection acting on other loci. If the distribution of genotypes in the seeds is known before sowing, *viability selection* can be detected by statistical testing as rejection of the null hypothesis that the survivors are a random sample of the sown seeds and thus the observed deviation of the distribution of genotypes in the survivors from the distribution in the seeds is an effect of random sampling. The practical problem is that the distribution among the seeds is rarely known, since genotyping destroys seeds in most species.

In many studies, samples of seeds are retained as reference samples, enabling a test of homogeneity between the reference sample and the survivors (i.e., a test of the hypothesis that the two sets are independent samples from the same base population of seeds) (Hufford and Hamrick, 2003). Reference samples are, however, often too small to be representative or even to contain all alleles or genotypes with sufficient probability. As a consequence, it can be difficult to test for viability selection directly.

Alternatively, the operation of viability selection at a locus can be concluded indirectly from the observation of statistically significant differences between environments for the distribution of genotypes among survivors. The prerequisite is that the seeds initially sown in each environment were random samples from the same seed lot. The occurrence of such *differential viability selection* implies that viability selection must have occurred in at least one of the environments.

So-called outlier methods based on F_{ST} (or G_{ST}) are commonly applied to detect gene loci that show significant differences among survivors between environments. These methods regard especially high or low values of F_{ST} as an indicator of directional or balancing selection, respectively (see review of Nosil et al., 2009). Commonly used software includes DFDIST (Beaumont and Nichols, 1996), LOSITAN (Antao et al., 2008), BAYESCAN (Foll and Gaggiotti, 2008). The performance of these methods has been analyzed in simulation studies (Lotterhos and Whitlock, 2014; Flanagan and Jones, 2015).

The index F_{ST} is, however, primarily a measure of fixation (Gregorius and Roberds, 1986). In the words of Wright (Wright, 1978, p. 82), the "fixation index [F_{ST}] is thus not a measure of degree of differentiation in the sense implied in the extreme case by absence of any common allele. It measures differentiation within the total array in the sense of the extent to which the process of fixation has gone toward completion" in the context of the Island Model. Because selection does not always favor only one allele in each environment, it is not a suitable measure of differential selection.

Another major disadvantage of the index F_{ST} is that it considers only the distributions of alleles. Each level of integration has its own set of genetic types, such as genes in a gene pool, haplotypes, single-locus genotypes or multi-locus genotypes (Gillet and Gregorius, 2008). The survivors in the different environments can differ for their distributions of genetic types at each such level of genetic integration. Differences can be significant for the distributions of genetic types at one level of genetic integration but not another. Thus F_{ST} outlier methods may

fail to detect selection at the genotype level, as is demonstrated below for an enzyme locus.

The measures of compositional differentiation δ_{SD} and Δ_{SD}, in contrast, directly measure the frequency differences between environments at any given level of genetic integration (Gregorius and Roberds, 1986; Gregorius et al., 2003; Gillet et al., 2004; Gillet and Gregorius, 2008). At still higher levels of genetic integration, multiple loci can be considered on the basis of measures of genic difference between multi-locus genotypes. This allows assessment of the joint effects of homologous and non-homologous (between loci) gene association on differential selection.

In this paper, we present a non-parametric permutation method for detecting differential viability selection among environments at any level of genetic integration. The method is applicable to an experimental design involving seeds harvested in one location and randomly sown in different environments. At a given level of genetic integration for specified loci, a significant P-value of the observed difference δ_{SD} or Δ_{SD} between survivors in the various environments is an indication that differential viability selection is acting at this level of genetic integration. Inferred differential selection can be due to direct functional selection on the locus or to stochastic association in the seeds between this locus and a second locus that is subject to direct functional selection. Calculations can be performed by the computer program *DifferInt* (Gillet, 2013). In a case study, the method is demonstrated using published data on enzyme gene loci obtained as part of a large-scale translocation experiment in beech (Müller-Starck, 1993). Results are plotted in diagrams that are reminiscent of the outlier methods but illustrate increasing levels of genetic integration.

Material and Methods

Hierarchical sampling design
Assume that the survivors in all environments stem from a single large sample of seeds that were harvested in a common environment, preferably in the same stand, and were of unknown genotypic frequency distribution for a specified set of gene loci. Sampling step 1: The seeds in this original sample were randomly divided into a finite number of samples of given sizes, and each of these samples was sown in a different environment. Due to the large size of the original seed sample, the (unknown) genotype frequencies are assumed to be equally distributed over the environments. Sampling step 2: After a specified period of time during which mortality could occur, consider the survivors in each environment as a sample drawn without replacement from the seeds that were sown there, and score their genotypes at specified gene loci.

Levels of genetic integration
The organization of genetic information in individuals is intrinsically hierarchical. At the lowest level, the gene pool as a "beanbag" of genes is organized into alleles at gene loci. The

alleles at each gene locus are integrated into the individual's genotype at the locus. Its genotypes at several loci are integrated into its multi-locus genotype at these loci, etc. The steps of this hierarchy are called *levels of genetic integration* by Gillet and Gregorius, (2008). For convenience, the combinations of genes (alleles) of a set of loci that can be formed at each integration level will be referred to as the genetic types at this level (i.e., genes/alleles, single- or multi-locus genotypes).

For one or more specified gene loci, the frequency distribution of genetic types can be determined at each integration level within a deme (i.e., a collection of individuals) by counting the number of occurrences of each genetic type and dividing it by the numer of individuals. At the gene pool level, the relative frequency of each gene equals its number of occurrences divided by the product of the number of loci, the degree of ploidy, and the number of individuals. At the single-locus level, the frequency of each genotype in the "genotype bag" is the number of occurrences divided by the product of the number of loci and the number of individuals. At the multi-locus level, the frequency of each multi-locus genotype is simply the number of individuals that possess it divided by the total number of individuals. Note that this way of counting occurrences considers two genetic types at the same integration level to be either the same or, if they differ by at least one allele at one of the loci, completely different. More gradual measurement of the difference between two genetic types in terms of the proportion of shared alleles will be considered below.

Measures

Genetic distance d_0

The measures of compositional differentiation applied here are based on the absolute genetic distance d_0 (Gregorius, 1974, 1984) between two demes at a given level of genetic integration for a given set of gene loci (see Table 1 for definitions of this and all subsequent measures). d_0 is applicable to any frequency distribution and can thus be calculated at any level of genetic integration. d_0 equals 0 if the frequency distributions are identical and 1 if they are disjoint, i.e., no individuals share the same type.

Complementary compositional differentiation δ_{SD}

For two or more demes, the *complementary compositional differentiation* δ_{SD} among demes at a given level of genetic integration for a given set of gene loci equals the mean d_0-distance of each deme to its complement, i.e., the deme formed by pooling the other demes (Gregorius and Roberds, 1986) (subscript SD for "symmetric distance"). The minimum of δ_{SD} =0 holds only when the frequency distributions are identical over all demes, the maximum of δ_{SD} =1 only when no two demes have individuals that share the same genetic type. δ_{SD} is calculated by *DifferInt* (Gillet, 2013).

Table 1
Notation and measures

D_j	Deme D_j ($j = 1, …, m$) contains genetic types i ($i = 1, …, n$) with relative frequency $p_i(j)$, such that $p_i(j) \geq 0$ and $\sum_{i=1}^n p_i(j) = 1$
c_j	The *weight* given to D_j, such that $c_j \geq 0$ and $\sum_{j=1}^m c_j = 1$. Usually $c_j \equiv 1/m$ or is proportional to deme size
D_j^c	*Complement deme* D_j^c is a pooling of all demes other than D_j. D_j^c contains genetic type i with relative frequency $\bar{p}_i(j) = \sum_{k \neq j} c_j \cdot p_i(j)/(1 - c_j)$
d_0	Absolute genetic distance $$d_0(D_j, D_{j'}) = \frac{1}{2}\sum_{i=1}^n |p_i(j) - p_i(j')|$$ is the minimal change necessary in D_j to make it match $D_{j'}$.
δ_{SD}	Complementary compositional differentiation $$\delta_{SD}(D_1, …, D_m) = \sum_{j=1}^m c_j \cdot d_0(D_j, D_j^c)$$
d	Genic difference $d(a, b)$ between genetic types a and b equals the proportion of genes (alleles) that a and b do not share
Δ	Genetic distance considering genic difference $$\Delta(D_j, D_{j'}) = \min \sum_{i=1}^n d(a, b) \cdot s(a, b)$$ minimized over all shift transformations that shift the amount $s(a, b)$ of each genetic type a that is more frequent in D_j than $D_{j'}$, to a genetic type b that is less frequent in D_j than $D_{j'}$, weighted by the genic difference $d(a, b)$. Shifting continues until the frequency distribution of types in D_j matches $D_{j'}$.
Δ_{SD}	Complementary compositional differentiation considering genic difference $$\Delta_{SD}(D_1, …, D_m) = \sum_{j=1}^m c_j \cdot \Delta(D_j, D_j^c)$$

Since each genetic type is counted as being qualitatively completely different from all others, d_0 and thus also δ_{SD} do not consider that two genetic types that share some of the same alleles differ less than two types that share no alleles. Thus the two-locus genotypes $A_1 A_1 B_1 B_1$ and $A_1 A_2 B_1 B_2$ that share 50 % of their alleles are just as completely different as the genotypes $A_1 A_1 B_1 B_1$ and $A_2 A_2 B_2 B_2$ that share no alleles.

Genic difference d

Define the *genic difference* d between two genetic types at a given integration level for a set of loci as the mean proportion of their alleles that they do not share at each locus. At the genic level for a single locus, for example, $d(A_i, A_i) = 0$ and $d(A_i, A_j) = 1$ hold for $i \neq j$. At the genotype level for a single locus, $d(A_i A_j, A_i A_i) = d(A_i A_j, A_i A_j) = 0$, $d(A_i A_j, A_i A_k) = 0.5$, and $d(A_i A_j, A_j A_j) = d(A_i A_j, A_j A_k) = d(A_i A_j, A_k A_l) = 1$ hold, where all indices are unequal.

Genetic distance Δ considering genic difference

When types are characterized by more than a single feature, they can be considered to be structured. Accordingly, types can be distinguished either without or with respect to structural characteristics. In the former, they are distinguished in a binary way by considering them to be different if they differ by at least one feature without reference to the number of features by which they differ. Thus for a single diploid locus, for example, two genotypes in which the alleles figure as features can be distinguished in two ways: either solely on the basis of their differing for at least one allele or for the number of alleles they do not share. In the second case, structural characteristics are considered, while in the first case they are not.

The genetic distance Δ between two demes additionally takes such structural differences between genetic types in the form of the above introduced genic differences into account (Gregorius et al., 2003; Gillet et al., 2004; Gillet and Gregorius, 2008). Δ equals the minimum "cost" of shifting genetic types in one of the demes to other genetic types in order to make the distribution of genetic types in this deme match the distribution in the other deme. In effect, individuals of a genetic type a that is in excess in the one deme are shifted to individuals of a genetic type b that is deficient in this deme, both in comparison to the frequencies of these types in the other deme. The cost of shifting one genetic type to another is the product of the relative frequency $s(a,b)$ of the shifted individuals and the genic difference $d(a,b)$ between the types. The "cost" of a shift transformation equals $\sum d(a,b) \cdot s(a,b)$ over all shifted types. Δ then equals the minimum cost over all possible shift transformations of the one deme to the other. Δ=0 holds when the distributions of genetic types do not differ between demes, i.e., the distributions in the demes already match. Δ=1 holds only when no allele at any of the loci is shared by the two demes, so that the genic difference for all shifts equals 1.

Complementary compositional differentiation Δ_{SD} considering genic distance

The measure Δ_{SD} of *complementary compositional differentiation considering genic difference* is analogous to δ_{SD}, except that the genetic distance of each deme to its complement is measured by Δ instead of d_0 (Gillet and Gregorius, 2008; Gillet, 2013). The minimum of $\Delta_{SD} = 0$ is assumed only when the frequency distributions are identical in all demes, and the maximum of $\Delta_{SD} = 1$ is reached only when no two demes have individuals that share the same allele. Δ_{SD} is calculated by *DifferInt* (Gillet, 2013).

Increment of differentiation due to structural differences

At any given integration level, Δ_{SD} cannot be less than δ_{SD} at any lower level, and their equality indicates that the pattern of association of the genetic types at the lower level into the genetic types at the given level follows the same pattern in all demes (Gillet and Gregorius, 2008). For a single locus, equality of δ_{SD} at the allelic level and Δ_{SD} at the genotype level indicates that the pattern of homologous gene association (i.e., the pattern by which alleles were combined into genotypes) is the same in all demes. In this case, differentiation as measured by Δ_{SD} is completely explained by allelic differentiation at this locus as measured by δ_{SD}. A positive value of $\Delta_{SD} - \delta_{SD}$, in contrast, indicates that structural characteristics such as homologous gene associations are involved in bringing about the genetic differentiation between environments. Allelic differences alone would not suffice to explain the differentiation.

It follows that for a single locus, the expression

$$I_{str} = \frac{\Delta_{SD}^{genotypes} - \delta_{SD}^{genes}}{\Delta_{SD}^{genotypes}}$$

measures the relative gain in differentiation due to differences in structural characteristics (patterns of gene association) among environments. At the one extreme, $I_{str} = 0$ holds when the genetic differentiation $\Delta_{SD}^{genotypes}$ is due solely to differences in allele frequencies at the loci (gene pool). At the other extreme, $I_{str} = 1$ holds when $\delta_{SD}^{genes} = 0$ and $\Delta_{SD}^{genotypes} > 0$, meaning that differences between demes are due only to differences in the pattern of gene association. I_{str} is not defined when its denominator $\Delta_{SD}^{genotypes}$ equals 0, but it makes no sense to calculate I_{str} in that case, since it follows from the inequality $\Delta_{SD}^{genotypes} \geq \delta_{SD}^{genes}$ that $\delta_{SD}^{genes} = 0$ also holds (Gillet and Gregorius, 2008).

Simpson diversity

The Simpson diversity for the alleles at a locus measures the probability

$$S = 1 - \sum_i p_i^2$$

of randomly drawing two different alleles (with replacement) from the beanbag of alleles, where p_i is the relative frequency of the i-th allele in the deme (Simpson, 1949). This measure is often called H_e, or "expected heterozygosity" among genotypes, but this name makes no sense unless the genotypes are expected to show Hardy-Weinberg-Proportions (HWP), such as when the deme is a sample of the progeny from random mating among a collection of individuals. The effective number of alleles at the locus equals $1/(1-S)$ (Gregorius, 1991). The Simpson diversity in the pool of all survivors is denoted by S_T (T for "total").

Fixation index F_{ST}

The most commonly used measures of genetic variation among demes are the diversity-oriented measures $F_{ST} = G_{ST}$ (Nei, 1973; Wright, 1978) and their relatives, including G_{ST}' (Hedrick, 2005). $F_{ST} = (H_T - H_S)/H_T$, for example, measures the differences at the allelic level between H_T, which equals the Simpson diversity $S_T = 1 - \sum_i p_i^2$ within the entire collection, and H_S, which is the mean of the Simpson diversities within the demes. Through division by H_T, these measures reach their maximum when $H_S = 0$, i.e., when all demes are monomorphic (i.e., no variation within demes), even if all but one of the demes are monomorphic for the same genetic type. They are never maximal if at least one deme is polymorphic, even if the demes are disjoint. An exception is G_{ST}, which is maximal in both cases: when all demes are monomorphic or when all demes are disjoint or both (Gregorius et al., 2007). Thus F_{ST} and related measures do not measure compositional differentiation (Gregorius and Roberds, 1986; Gregorius et al., 2007; Jost, 2008). Jost's D (Jost, 2008), which is also a standardized measure of the difference between H_T and H_S at the allelic level, assumes its maximum of 1 only for complete compositional differentiation, i.e.,

when demes share no alleles. None of these measures are applied to higher levels of genetic integration.

Homozygote excess

Under the expectation that the genotypes at a locus should show HWP, the relative excess of homozygotes can be expressed as

$$F_{IS} = 1 - H_o/H_e$$

where $H_e = 1 - \Sigma_i\, p_i^2 = S$ is the expected proportion of heterozygotes under HWP and H_o is the observed proportion of heterozygotes.

Testing for differential viability selection

When analyzing real sets of survivors in different environments, consider the genes at one or more gene loci and the genetic types that occur at a given integration level. Assume that the seeds sown in each environment were a large random sample from the same infinitely large seed lot, so that the genetic types in the seeds were identically distributed in all environments. Subsequent selection may change the distribution of genetic types in the environments, but if the selection does not occur differentially between environments, then the distribution of genetic types among the survivors should also be the same in all environments. This means that any observed differences between environments for the distribution of the genetic types among the survivors would be due to effects of random sampling of the relatively small number of survivors. In statistical terms, the occurrence of differential viability selection is indicated by rejection of the following

Null hypothesis: The survivors in each environment are a random sample from the same base population.

In principle, all tests of the null hypothesis base their conclusions on some measure of the discrepancy between the actually observed distributions. The probability of obtaining a discrepancy that is at least as large as the observed discrepancy if the hypothesis is true is termed the P-value for this measure. The hypothesis is rejected if the P-value is small, usually <0.05.

Classical tests of homogeneity

The Pearson Chi-square and the Likelihood-ratio tests of homogeneity of distributions are classical parametric tests of the hypothesis that the individuals found in each of a set of populations are random samples from the same (or identically distributed) base population(s). If individuals are classified by the states (types) of a given trait, these homogeneity tests are conditional on marginal distribution of the types among all sampled individuals and the given sample sizes. Conditionality means that the test considers only random samples that fulfill these conditions. For random sampling with replacement (or without replacement from a very large base population), the marginal distribution of types happens to be the maximum

likelihood estimate (MLE) of this distribution. For the Pearson Chi-square test, the measure of discrepancy among the distributions observed in the populations is the minimum-variance measure X^2; for the Likelihood-ratio test, it is the log-likelihood-ratio G. The P-value of the observed discrepancy approaches $1-(\mathcal{X}_n^2)^{-1}$ (X^2) or $1-(\mathcal{X}_n^2)^{-1}$ (G) asymptotically (i.e., for increasingly large samples), where $(\mathcal{X}_n^2)^{-1}$ is the inverse function of the $\mathcal{X}_n^2$-function for n degrees of freedom. The hypothesis is rejected if the P-value is less than the chosen level of significance (usually 0.05). The estimates are often considered to be sufficiently exact if the expected number of individuals of any type in any sample is at least 5, since otherwise the approximate P-value can be "artificially" small, leading to false rejection of the hypothesis. Other assumptions concerning the distributions of $\mathcal{X}^2$ and G require verification.

Interpretation: The hypothesis of the absence of differential selection at a specified integration level for given loci is rejected if the P-value is less than the chosen level of significance (usually 0.05, 0.01, or 0.001).

- Significance for alleles at a locus implies differential allelic selection among the environments.
- Significance for the genotypes at a single locus implies differential genotypic selection among the environments.
- Significance for multi-locus genotypes implies differential selection of multi-locus genotypes among the environments.

These classical tests are not tests of gene association. They are based on type frequencies and their differences between collections. They do not consider variable differences between types as does Δ_{SD}. Thus, with classical tests one could obtain large significance irrespective of the amount of genic association in the genotypes.

Permutation test of the absence of differential viability selection

The null hypothesis describes the case where differences in the viability of seeds may depend on differences in their genetic types but not on differences between the environments in which they are sown. Thus all seeds of the same genetic type have the same viability in all environments. We do not know the exact distribution of genetic types in the seeds, but rather only that it was the same in all environments. Thus we cannot simulate random sampling directly. We can, however, randomly reassign the actual survivors to environments in the numbers in which they were observed, and every particular assignment would have the same probability of occurring. Each such assignment corresponds to a permutation of the actual survivors among the environments.

In a permutation test, random assignment can be simulated by generating all permutations of the survivors among the environments or, when the total number of permutations is computationally prohibitive, by simulating a random sample of permutations. For each permutation, a given measure of the discrepancy among the distributions of genetic types among

environments is calculated as the characteristic value of the permutation. The P-value then equals the proportion of permutations that yield a difference that is greater than or equal to the observed difference. The hypothesis of random assignment is rejected if the P-value is small (usually <0.05). Rejection, in turn, is an indication that differential viability selection occurred among environments. Calculations are performed using an extended version of *DifferInt* (Gillet, 2013).

While Δ_{SD} and δ_{SD} are measures of differentiation, they also can be interpreted as measures of association of genetic types with environments (Gregorius, 2011). Thus the absence of differentiation ($\delta_{SD} = \Delta_{SD} = 0$), where genetic type distributions are identical in all environments, implies the absence of (or independent) association between genetic type and environment. Complete differentiation without consideration of structural characteristics ($\delta_{SD} = 1$), where no genetic types occurs in more than one environment, implies complete association of genetic type with environment. Complete differentiation with structural consideration ($\Delta_{SD} = 1$), where no allele occurs in more than one environment, implies complete association of the alleles at each locus with the environment. As such, these measures are suitable measures of discrepancy for permutation testing of the null hypothesis of the absence of association between the genetic type of a survivor and the environment in which it is found. If the hypothesis is true, any observed differentiation would be due to random effects in the assignment of survivors to environments without regard for their genetic types.

Interpretation: A significantly small Δ_{SD} or δ_{SD} (P-value<0.05) leads to rejection of the null hypothesis that the viability of all genetic types is independent of the environment at the respective integration level and loci in question.

- Significance of δ_{SD} at the gene pool level implies differential selection of alleles among the environments.
- Significance of Δ_{SD} at the single-locus genotype level implies differential selection of genotypes among the environments when genic differences and structural differences for genic association patterns are considered
- Significance of δ_{SD} at the genotype level implies differential selection of genotypes among the environments without consideration of genetic structure (genic differences between genotypes).
- Significance of Δ_{SD} at the multi-locus genotype level implies differential selection of multi-locus genotypes among environments when genic differences and structural differences for genic association patterns are considered.
- Significance of δ_{SD} at the multi-locus genotype level implies differential selection of multi-locus genotypes among the environments without consideration of genetic structure (genic differences between genotypes).

The main difference between δ_{SD} and Δ_{SD} is that δ_{SD} does not take "structural" (or variable) differences between genotypes into account. δ_{SD} treats two genotypes (as well as two multi-locus haplotypes) as being either the same or completely different, just as if they were alleles at one locus. Permutation tests could also be performed using X^2 or G as discrepancy measures that do not consider variable differences between genotypes (without regard for their asymptotic X^2-distribution), but they lack the elementary properties of differentiation measures.

Permutation test of the contribution of the structural increment to genotypic differentiation

Using the same permutations, the hypothesis can be tested that the structural increment I_{str} of genotypic differentiation, i.e., differences in the patterns of homologous gene association over environments, has a significant effect on genotypic differentiation. The P-value of the structural increment I_{str} over the permutations equals the proportion of permutations that yield a value of I_{str} greater than or equal to the value of I_{str} observed among the survivors.

Interpretation: Both large and small P-values of I_{str} have implications:

- A significantly large I_{str} (P-value>0.95) implies that the differences in patterns of gene association are much smaller than expected for random assignment of individuals to environments, i.e., differential selection for the pattern of homologous gene association is very weak.
- A significantly small I_{str} (P-value<0.05) implies that the differences in patterns of gene association are much larger than expected for random assignment of individuals to environments, i.e., differential selection for patterns of gene association is very strong.

Case study

In a large field study in European beech (*Fagus sylvatica* L.), seeds were harvested in the year 1984 at several locations (Müller-Starck, 1993). For each seed source, a random sample of seeds were pre-germinated, and the germinated seeds were randomly divided into smaller samples that were sown in different environments. Here we consider one seed source (St. Märgen, Black Forest) and three environments. A sample of 1500 germinated seeds was sown in environment *Arnsberg* (Sauerland), a forest plot with acidic soil. A second sample of 260 germinated seeds was sown in *Nursery*, a plot of standard garden soil close to the department's greenhouse in Göttingen. A third sample of 1500 germinated seeds was sown in *Bramwald*, another forest plot with acidic soil. After two years, the surviving juvenile trees were counted and scored for their genotypes at twelve enzyme gene loci that were thought to have selective potential. Mortality was moderate in *Nursery*, high in *Arnsberg* and extremely high in *Bramwald* (Table 2). Our demonstrations concentrate on the two environments *Arnsberg* and *Nursery* because of their lower mortality and large environmental difference. The extreme mortality in *Bramwald* suggests that the environmental conditions were unsuitable at the outset for the seed sample and might therefore have produced unrealistic

Table 2

Source location of the seeds and the survival rates (No. of survivors / No. of sown seeds) in three environments.

Seed Source		Environments and survival rate		
Location St. Märgen	Altitude m.a.s.l.	*Arnsberg* (acidic soil)	*Nursery* (garden soil)	*Bramwald* (acidic soil)
	895	0.153 (229/1500)	0.738 (192/260)	0.035 (53/1500)

m.a.s.l. - meters above sea level

Table 3

Absolute frequencies of alleles and genotypes at all twelve enzyme gene loci in the field study (Müller-Starck, 1993). Allele A_i is designated by i, genotype $A_i A_j$ by ij. Loci are ordered by ascending Simpson diversity S_T in the two environments *Arnsberg* and *Nursery*. *Bramwald* is excluded from main calculations due to the extreme mortality

Locus	Alleles					Genotypes								
PGI-B	**2**	**3**				**22**	**33**	**23**						
Arnsberg	457	1				228	0	1						
Nursery	381	3				190	1	1						
Bramwald	105	1				52	0	1						
DIA-A	**2**	**3**				**22**	**33**	**23**						
Arnsberg	450	8				222	1	6						
Nursery	376	8				184	0	8						
Bramwald	106	0				53	0	0						
SKDH-A	**2**	**3**	**4**	**5**		**33**	**55**	**23**	**34**	**35**				
Arnsberg	12	433	0	11		207	2	12	0	7				
Nursery	0	377	1	6		187	2	0	1	2				
Bramwald	0	103	0	3		50	0	0	0	3				
6PGH-B	**2**	**3**	**4**			**22**	**33**	**23**	**24**	**34**				
Arnsberg	29	426	3			0	198	28	1	2				
Nursery	16	362	6			2	172	12	0	6				
Bramwald	1	105	0			0	52	1	0	0				
6PGH-A	**2**	**3**				**22**	**33**	**23**						
Arnsberg	403	55				176	2	51						
Nursery	336	48				147	3	42						
Bramwald	97	9				45	1	7						
MDH-B	**2**	**3**	**4**	**5**	**6**	**22**	**33**	**44**	**55**	**66**	**23**	**24**	**25**	**35**
Arnsberg	31	37	13	375	2	2	1	6	158	1	1	1	25	34
Nursery	20	36	5	319	4	1	6	2	139	2	0	1	17	24
Bramwald	8	7	2	89	0	2	1	1	40	0	0	0	4	5
MDH-C	**2**	**4**				**22**	**44**	**24**						
Arnsberg	135	323				21	115	93						
Nursery	119	265				23	96	73						
Bramwald	26	80				4	31	18						
PGM-A	**2**	**3**	**4**			**33**	**44**	**23**	**34**					
Arnsberg	2	132	324			29	126	2	72					
Nursery	0	122	262			28	98	0	66					
Bramwald	0	23	83			5	35	0	13					
IDH-A	**2**	**3**	**4**			**22**	**33**	**23**	**24**	**34**				
Arnsberg	149	307	2			20	99	108	1	1				
Nursery	122	261	1			17	86	88	0	1				
Bramwald	36	70	0			7	24	22	0	0				
PER-B	**1**	**2**	**3**	**4**	**6**	**33**	**44**	**66**	**14**	**23**	**34**	**36**	**46**	
Arnsberg	1	3	161	277	16	48	102	0	1	3	59	3	13	
Nursery	0	0	148	231	5	46	86	1	0	0	56	0	3	
Bramwald	0	0	23	78	5	2	28	0	0	0	18	1	4	

AAT-B	1	2	3	4	22	33	12	23	34				
Arnsberg	0	187	268	3	33	72	0	121	3				
Nursery	15	159	203	7	41	67	15	62	7				
Bramwald	0	39	58	9	8	13	0	23	9				
LAP-A	1	2	3	4	22	33	44	12	13	14	23	24	34
Arnsberg	3	139	198	118	39	67	38	1	1	1	41	19	22
Nursery	3	162	181	38	66	71	13	0	2	1	28	2	9
Bramwald	1	31	36	38	5	10	12	1	0	0	11	9	5

results. The absolute distributions of alleles and genotypes at all twelve loci are listed in Table 3.

No differential selection

For nine of the twelve loci, testing based on a (pseudo-)random sample of 5000 permutations found no indication of significant differential selection between *Arnsberg* and *Nursery* at any integration level, when the three measures of compositional differentiation were used as measures of discrepancy (i.e., all P-values are larger than 0.05). These loci are PGI-B, DIA-A, 6PGH-B and -C, MDH-B and -C, PGM-A, IDH-A, and PER-B (see Table 4 and Figure 1, bottom graph). The structural increment I_{str} and the fixation index F_{ST} are also non-significant for these loci. It can therefore be concluded that selection acted on the genotype at each of these loci independently of the environment (though this does not exclude the possibility that differential selection occurred jointly at multiple loci).

Differential selection for genotypes

The disregard of the complementary compositional differentiation $\delta_{SD}^{genotypes}$ for any effects of differential selection between environments on the distributions of the alleles and their association among survivors lends it the role of an unstructured backdrop for the structural analysis in subsequent sections.

At three loci, the same 5000 permutations yielded highly significant discrepancy between *Arnsberg* and *Nursery* for $\delta_{SD}^{genotypes}$ (P-value<0.01). One is the nearly monomorphic locus SKDH-A (shikimate 5-dehydrogenase, EC 1.1.1.25) with total diversity S_T=0.070. The other two happen to be the two most polymorphic loci: AAT-B (aspartate aminotransferase, EC 2.6.1.1) with two frequent alleles and two rare ones (S_T=0.518) and LAP-A (leucine aminopeptidase EC 3.4.11.1) with three frequent alleles and one rare one (S_T=0.635) (see distributions in Table 3). These three loci also show the highest compositional differentiation $\delta_{SD}^{genotypes}$ between environments among the twelve loci. This implies differential selection between environments for the genotypes at these three loci, either directly or due to stochastic association with functional loci.

Differential selection for alleles

Of the three loci that show differential selection for genotypes by $\delta_{SD}^{genotypes}$, two also show significance for δ_{SD}^{genes}.

• SKDH-A: δ_{SD}^{genes} = 0.035* is small but significant for this nearly monomorphic locus. Note that the proportion of the three rare alleles is higher in *Arnsberg* (23/456=0.050) than in *Nursery* (7/384=0.018), and the rare allele A_2 that occurs in the appreciable proportion of 0.026 in *Arnsberg* is even absent in *Nursery*. This is surprising considering the low mortality of only 0.262 in the benign environment *Nursery* compared to 0.847 in *Arnsberg*, since if mortality had been random, more of the initial variation in the seeds should have survived in *Nursery* than in *Arnsberg*. This suggests that possession of a rare allele, especially A_2, may be detrimental for survival under benign conditions.

• LAP-A: δ_{SD}^{genes} = 0.159*** is by far the largest over all loci and the most highly significant. Allele A_4 is much more frequent following the high mortality in *Arnsberg* (0.258) than the low mortality in *Nursery* (0.099), suggesting that A_4 must be very beneficial for survival in the harsh environment or detrimental under benign conditions.

• AAT-B: δ_{SD}^{genes} = 0.057 is not significant. It is worth noting that the allele B_1 that survives at frequency 0.039 in *Nursery* does not survive the high mortality of *Arnsberg*.

This implies that the individual alleles at SKDH-A and LAP-A have an influence on survival that differs between the two environments, either directly or stochastically. The lack of significance for AAT-B indicates that the differential selection on the "unstructured" genotypes ($\delta_{SD}^{genotypes}$) did not affect specific alleles at this locus.

Differential selection of genotypes considering genic differences

All three of the loci that show significant compositional differentiation $\delta_{SD}^{genotypes}$ also show significance when structural differences are considered in $\Delta_{SD}^{genotypes}$. This means that consideration of genic differences between genotypes reveals that the pattern by which the alleles are associated in the genotypes, which is hypothesized to have been the same in the sown seeds, has significantly different effects on survival in the two environments. Two of these loci also show significance for δ_{SD}^{genes}, indicating that differences in the pattern of gene association do not conceal the effects of selective differences at the lower level of the alleles. For the third locus, AAT-B, no differential selection at the allelic level is apparent. Interpretation of the structural increment I_{str} can show whether differential selection was stronger on the alleles or on their patterns of association.
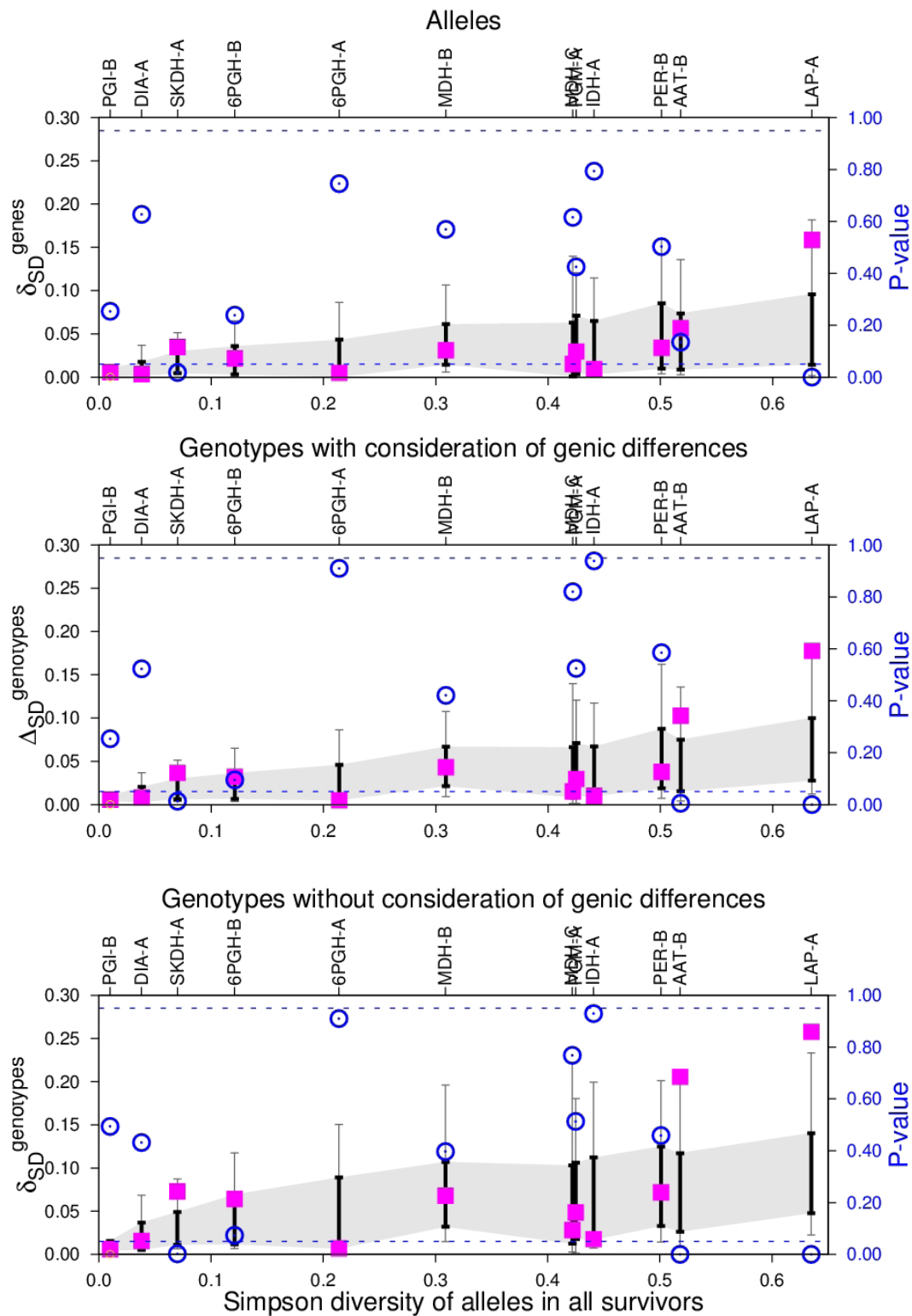
Figure 1

Compositional differentiation between survivors in *Arnsberg* and in *Nursery* at the different genetic integration levels. Loci are arranged on the *x*-axis by allelic diversity S_T among all survivors (MDH-C is slightly less diverse than PGM-A). For each locus, the magenta-colored box marks the observed compositional differentiation, the thick black line covers the central 90 % of the differentiation values for all permutations, and the thin black line spans the range of all permutations. Open circles mark the P-values on the scale of the axis to the right. For LAP-A and SKDH-A, differential selection is significant at all levels (P-value<0.05). For AAT-B, differential selection is significant for genotypes with and without consideration of genic differences but not for alleles.

Table 4

Two environments *Arnsberg* and *Nursery* : Results of 5000 random permutations of survivors between environments. Loci are ordered by increasing total Simpson diversity S_T for alleles among survivors. The observed value of each measure is listed for each locus, followed by the P-value in square brackets. Asterisks denote level of significance of P-values as <0.05*, <0.01**, and <0.001*** (0.000 refers to a P-value less than 0.001)

			Results for two environments Arnsberg and Nursery			
Locus	S_T	δ_{SD}^{genes}	$\Delta_{SD}^{genotypes}$	$\delta_{SD}^{genotypes}$	I_{str}	F_{ST}
PGI-B	0.010	0.006 [0.254]	0.006 [0.254]	0.006 [0.494]	0.000 [0.240]	0.002 [0.254]
DIA-A	0.038	0.003 [0.627]	0.008 [0.523]	0.016 [0.432]	0.565 [0.079]	0.000 [0.627]
SKDH-A	0.070	0.035 [0.018 *]	0.037 [0.013 *]	0.073 [0.002 **]	0.045 [0.400]	0.007 [0.019 *]
6PGH-B	0.121	0.022 [0.239]	0.032 [0.095]	0.064 [0.073]	0.325 [0.213]	0.002 [0.289]
6PGH-A	0.215	0.005 [0.746]	0.005 [0.910]	0.007 [0.910]	0.000 [0.738]	0.000 [0.746]
MDH-B	0.309	0.031 [0.570]	0.043 [0.421]	0.068 [0.397]	0.282 [0.226]	0.001 [0.671]
MDH-C	0.422	0.015 [0.616]	0.015 [0.820]	0.028 [0.768]	0.000 [0.972]	0.000 [0.616]
PGM-A	0.425	0.030 [0.425]	0.030 [0.525]	0.049 [0.513]	0.000 [0.543]	0.001 [0.444]
IDH-A	0.441	0.009 [0.794]	0.010 [0.939]	0.018 [0.929]	0.060 [0.474]	0.000 [0.774]
PER-B	0.501	0.034 [0.503]	0.038 [0.585]	0.072 [0.459]	0.102 [0.440]	0.001 [0.499]
AAT-B	0.518	0.057 [0.135]	0.103 [0.005 **]	0.206 [0.000 ***]	0.450 [0.145]	0.002 [0.153]
LAP-A	0.635	0.159 [0.000 ***]	0.178 [0.000 ***]	0.258 [0.000 ***]	0.106 [0.528]	0.016 [0.000 ***]

Structural characteristics of genotypic differentiation

As mentioned above, genotypes can be differentially selected by their single alleles at a locus or by the association of the alleles in genotypes. Comparison of δ_{SD}^{genes} and the structural increment I_{str} can show which effect was more pronounced (see Table 4).

- SKDH-A: The near monomorphism at this locus forces genotypic differentiation $\delta_{SD}^{genotypes} = 0.073^{**}$ and $\Delta_{SD}^{genotypes} = 0.037^*$ to be small, but both are significant. The significance of $\delta_{SD}^{genes} = 0.035^*$ and the small size of the structural increment $I_{str} = 0.045$ suggest that differential selection acted primarily on the alleles and not on differences in their patterns of association. Note that significance of I_{str} is not to be expected due to the necessarily small sizes of both $\delta_{SD}^{genotypes} = 0.073^{**}$ and $\Delta_{SD}^{genotypes} = 0.037^*$ and the resulting wide range of I_{str} between 0 and 0.838 over the permutations.
- LAP-A: More differentiation is possible at this polymorphic locus, and $\delta_{SD}^{genotypes} = 0.258^{***}$ and $\Delta_{SD}^{genotypes} = 0.178^{***}$ are highly significant. The significance of the allelic differentiation $\delta_{SD}^{genes} = 0.159^{***}$ and the moderate size of $I_{str} = 0.106$ lead to the conclusion that selection operates mainly on the alleles at this locus but that patterns of association also had an effect.
- AAT-B: Though both $\delta_{SD}^{genotypes} = 0.206^{***}$ and Δ_{SD} genotypes = 0.103^{**} are significant, the allelic differentiation $\delta_{SD}^{genes} = 0.057$ and the large structural increment $I_{str} = 0.450$ are not. It can be concluded that whereas differential selection on the alleles and their patterns of association is not significantly strong when viewed separately,

their interaction creates significant differential selection for genotypes.

Direct measurement of gene association by homozygote excess

Inbreeding is one type of gene association pattern, and it yields an excess of homozygotes relative to HWP, i.e., F_{IS} is greater than 0. Large differences between environments for F_{IS} should show up as a large value of I_{str}. It should be kept in mind that this measure does not cover all of the association patterns that can be detected by I_{str}.

- SKDH-A: The number of heterozygous survivors does not deviate significantly from the expectation under HWP in either environment (Log-likelihood Ratio Test). Thus the large difference in the amount of homozygote excess detected by F_{IS}, which in *Nursery* ($F_{IS} = 0.565$) amounts to four times that in *Arnsberg* ($F_{IS} = 0.141$), may be an artefact of the near monomorphism at this locus, or it may be an indication that homozygotes at this locus had slightly poorer chances of surviving in the harsh environment. The negligibility of the difference between association patterns is reflected by the small relative increment $I_{str} = 0.045$.
- LAP-A: In both environments, the number of heterozygous survivors shows highly significant deviation from the expectation under HWP (P<0.001 for both) at this polymorphic locus. The homozygote excess in *Nursery* ($F_{IS} = 0.629$) is greater than in *Arnsberg* ($F_{IS} = 0.433$), suggesting that homozygotes at this locus had poorer chances of surviving in the harsh environment. The difference in homozygote excess also indicates that association patterns play a role in the highly significant genotypic differentiation

between the environments, as reflected in the moderately positive relative increment $I_{str} = 0.106$.

- AAT-B: The number of heterozygous survivors in *Nursery* deviates significantly from the expectation under HWP (P=0.002) through an excess of homozygotes (F_{is}= 0.201), while the survivors in *Arnsberg* show a non-significant (P=0.125) excess of heterozygotes (F_{is}=-0.103). Homozygotes thus had considerably poorer chances of surviving in the harsh environment. The large relative increment $I_{str} = 0.450$ reflects this difference in association pattern.

Differential selection by fixation index

Recall that F_{ST} assumes small values both for nearly identical populations (small δ_{SD}^{genes}) and for highly polymorphic populations, regardless of their compositional differentiation δ_{SD}^{genes} (Hedrick, 2005; Gregorius et al., 2007). F_{ST} assumes large values only if all populations are nearly monomorphic for one of at least two alleles, and it is undefined if all populations are purely monomorphic for the same allele. The first three loci in Table 4 are almost monomorphic for the same allele (S_T<0.07), which poses the dilemma of whether F_{ST} should be large because of the small size of its denominator $S_T=H_T$ or small due to near identity. The observation that F_{ST} is very small and that its behavior parallels that of δ_{SD}^{genes} suggests that the differences between the populations have a greater influence on F_{ST} than the monomorphism. This becomes apparent from the fact that for low polymorphism populations are forced to share the same allele and have only negligible chances to differ for the rare alleles. Over the loci in the upper range of diversities, excepting LAP-A, F_{ST} remains exceedingly small, not reflecting the substantial variation in differentiation measured by δ_{SD}^{genes}. Now the rising polymorphism seems to determine the behavior of F_{ST}. Neither of these observations follows from inspection of F_{ST} only but rather requires additional consideration of the differentiation measures as well as the diversity measures.

- SKDH-A: The significance of δ_{SD}^{genes} for the nearly monomorphic locus SKDH-A means that these two environments produced nearly the maximum possible differential selection, and the significance of the very small F_{ST} shows that influence of the difference between populations dominates the vicinity to the point of non-definition.
- LAP-A: For the most highly polymorphic locus, both δ_{SD}^{genes} and F_{ST} have their largest values, and both are significant. Differential selection obviously created the highly significant difference between the populations, as measured by δ_{SD}^{genes}, but it is not possible to determine whether the high significance of F_{ST} is due to the large difference between populations or to a tendency towards fixation to different alleles. However, the allelic polymorphism apparently has yet not reached a level that is high enough to decisively lower F_{ST}.
- AAT-B: This locus provides a good demonstration that if only the non-significance of the fixation index F_{ST}=0.002

were considered, the significant differential selection revealed by $\Delta_{SD}^{genotypes}$ and $\delta_{SD}^{genotypes}$ at the higher integration levels would have been missed.

Multiple testing

If the same statistical test is performed on different data sets that were presumably obtained under the same conditions, and if the test hypothesis is true, a proportion of the tests equal to the significance level is nevertheless expected to yield rejection of the hypothesis. In this case, the hypothesis would be rejected if the P-value is less than more stringent significance levels that results by division of the "normal" levels (0.05[*], 0.01[**], 0.001[***]) by the number of multiple tests (Bonferroni correction). Regarding the twelve single-locus tests as twelve tests of the same locus (though the numbers of genetic types and their distributions certainly differed between loci in the original seed lot), significance levels are divided by 12. Single tests may lose their significance under the Bonferroni correction, as shown in Table 5. SKDH-A and AAT-B lose their significance for δ_{SD}^{genes} (AAT-B) and $\Delta_{SD}^{genotypes}$ but keep it for $\delta_{SD}^{genotypes}$. The finding that LAP-A retains the highest level of significance for all three differentiation measures even under multiple testing in combination with the moderate size of I_{str} points to a remarkably high degree of differential selection for the alleles at this locus.

Inclusion of a third environment

In the larger field study, seeds from a third random sample from the same seed lot (St. Märgen) were sown in a second forest environment *Bramwald*, and the survivors were scored at the same twelve enzyme gene loci after the same length of time. The ranking of the loci by the total diversity of alleles S_T did not change by adding the third environment, indicating that the allelic distributions in this environment were not too different from the means of the distributions in the other two environments.

Permutation tests of differential selection were performed by randomly permuting all survivors over the three environments. The measures and their P-values are shown in Table 6. All but three of the loci show the same pattern of (non-)significance at the three integration levels as for the two environments, including the highly polymorphic loci AAT-B and LAP-A. One of the exceptions is the nearly monomorphic locus SKDH-A, which loses its significance at all three integration levels. The other two are the nearly monomorphic locus 6PGH-B and the polymorphic locus PER-B, which gain significance at all levels.

Concluding remarks

The case study demonstrates the usefulness of the measures of compositional differentiation at different levels of genetic integration for the detection of differential selection between environments and for attributing selection to lower integration levels.

The three loci that show the significant differential selection on the genotype level are special, in that selection can be

Table 5

Multiple testing of differentiation between *Arnsberg* and *Nursery* (5000 permutations): P-values are listed for each locus. Asterisks denote significance for single testing (left) and for multiple testing of all twelve loci (right) using the Bonferroni correction (levels of significance divided by 12). SKDH-A becomes non-significant for δ_{SD}^{genes} and $\Delta_{SD}^{genotypes}$, and AAT-B loses significance for $\Delta_{SD}^{genotypes}$. $\delta_{SD}^{genotypes}$ retains significance at all three loci.

	Tests of single loci Levels of significance <0.05*, <0.01**, <0.001***			Multiple test of all twelve loci Levels of significance <0.00416*, <0.00083**, <0.00008***		
Locus	δ_{SD}^{genes}	$\Delta_{SD}^{genotypes}$	$\delta_{SD}^{genotypes}$	δ_{SD}^{genes}	$\Delta_{SD}^{genotypes}$	$\delta_{SD}^{genotypes}$
PGI-B						
DIA-A						
SKDH-A	*	*	**			*
6PGH-B						
6PGH-A						
MDH-B						
MDH-C						
PGM-A						
IDH-A						
PER-B						
AAT-B		**	**			**
LAP-A	***	***	***	**	***	***

Table 6

Three environments *Arnsberg*, *Nursery*, and *Bramwald*: Results of 5000 random permutations of survivors among environments, presented as for the two environments in Table 4. In comparison, significance is lost for SKDH-A but gained for 6PGH-B and PER-B.

			Results for three environments Arnsberg, Nursery, Bramwald			
Locus	S_T	δ_{SD}^{genes}	$\Delta_{SD}^{genotypes}$	$\delta_{SD}^{genotypes}$	I_{str}	F_{ST}
PGI-B	0.011	0.004 [0.514]	0.006 [0.451]	0.010 [0.362]	0.228 [0.197]	0.002 [0.514]
DIA-A	0.033	0.013 [0.192]	0.013 [0.215]	0.025 [0.173]	0.054 [0.785]	0.002 [0.341]
SKDH-A	0.068	0.022 [0.195]	0.026 [0.149]	0.051 [0.056]	0.140 [0.353]	0.006 [0.054]
6PGH-B	0.110	0.037 [0.030 *]	0.039 [0.031 *]	0.075 [0.024 *]	0.054 [0.687 *]	0.006 [0.049 *]
6PGH-A	0.208	0.025 [0.298]	0.030 [0.282]	0.060 [0.196]	0.168 [0.394]	0.001 [0.512]
MDH-B	0.307	0.026 [0.898]	0.041 [0.758]	0.067 [0.672]	0.367 [0.069]	0.001 [0.889]
MDH-C	0.416	0.038 [0.311]	0.040 [0.392]	0.058 [0.493]	0.049 [0.609]	0.002 [0.459]
PGM-A	0.417	0.060 [0.098]	0.060 [0.133]	0.088 [0.165]	0.000 [0.833]	0.004 [0.193]
IDH-A	0.442	0.012 [0.897]	0.021 [0.890]	0.040 [0.845]	0.432 [0.162]	0.000 [0.910]
PER-B	0.494	0.107 [0.011 *]	0.109 [0.012 *]	0.141 [0.027 *]	0.016 [0.877 *]	0.009 [0.035 *]
AAT-B	0.523	0.067 [0.096]	0.099 [0.011 *]	0.178 [0.001 ***]	0.318 [0.171]	0.004 [0.189]
LAP-A	0.644	0.149 [0.001 ***]	0.168 [0.000 ***]	0.244 [0.000 ***]	0.115 [0.586 ***]	0.021 [0.000 ***]

assigned to three combinations of differential selection on alleles or their patterns of association as measured by δ_{SD}^{genes} and I_{str}, respectively. At the nearly monomorphic locus SKDH-A, only the alleles are differentially selected between environments. At LAP-A, differential selection affects the alleles more than their patterns of association. At AAT-B, differential selection is not significant for alleles nor for their association patterns when viewed separately, but together they produce the significant differential selection for genotypes.

Commonly used methods for detecting selection would not have recognized the special characteristics of these loci:

- The problems with classical parametric tests of homogeneity using X^2 or the log-likelihood ratio G were mentioned above.

- Whereas outlier methods based on the fixation index F_{ST} would have reported differential selection at the allele level for SKDH-A and LAP-A, they would have missed the differential selection of genotypes for the locus AAT-B. This demonstrates that our method based on compositional differentiation at two integration levels is better able to detect differential selection than the F_{ST}-outlier methods.

- In the common analysis of the relative excess of homozygotes, gene association is measured by F_{IS} among the survivors within each environment separately. When F_{IS} differs between environments, there is no way to disentangle effects of the pattern of gene association and effects of differences in the underlying allelic distributions. The solution presented here is to divide the increment of genotypic over allelic differentiation into the contributions I_{str} of structural differentiation and (1- I_{str}) of allelic differentiation. The latter effect, when detected, can nevertheless be interpreted with the help of F_{IS}. For the three loci that showed differential selection, the association of like alleles (homozygotes) had better chances of surviving in the benign environment of *Nursery* than in the harsh forest environment of *Arnsberg*. Considering that mortality in *Arnsberg* was more than three times as high as the low mortality of 26% in *Nursery,* this result can be explained by a real selective disadvantage for homozygotes in *Arnsberg.*

References

Antao T, A Lopes, RJ Lopes, A Beja-Pereira and G Luikart (2008) LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. BMC Bioinformatics 9:323. https://dx.doi.org/10.1186/1471-2105-9-323

Beaumont MA and RA Nichols (1996) Evaluating loci for use in the genetic analysis of population structure. Proceedings of the Royal Society B: Biological Sciences 263(1377)1619-1626. https://dx.doi.org/10.1098/rspb.1996.0237

Flanagan SP and AG Jones (2015) Identifying signatures of sexual selection using genomewide selection components analysis. Ecology and Evolution 5(13):2722-2744. https://dx.doi.org/10.1002/ece3.1546

Foll M and O Gaggiotti (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. Genetics 180(2):977-993.

Gillet EM (2013) DifferInt : compositional differentiation among populations at three levels of genetic integration. Molecular Ecology Resources 13(5):953-964. https://dx.doi.org/10.1111/1755-0998.12145

Gillet EM and H-R Gregorius (2008) Measuring differentiation among populations at different levels of genetic integration. BMC Genetics 9:60.

Gillet EM, H-R Gregorius and M Ziehe (2004) May inclusion of trait differences in genetic cluster analysis alter our views? Forest Ecology and Management 197(1-3):149-158.

Gregorius H-R (1974) On the concept of genetic distance between populations based on gene frequencies. In: Department of Forest Genetics, Royal College of Forestry, Stockholm (ed) Proceedings of the IRUFRO Joint Meeting of Working Parties on Population and Ecological Genetics, Breeding Theory and Progeny Testing, S.02-04. Stockholm, pp 17-22. http://www.uni-goettingen.de/de/document/download/9afb2d3c2d8aee7036cc05d475cd5a64.pdf/017-Gregorius-IUFRO-1974.pdf

Gregorius H-R (1984) A unique genetic distance. Biometrical Journal 26(1):13-18. https://dx.doi.org/10.1002/bimj.4710260103

Gregorius H-R (2011) The analysis of association between traits when differences between trait states matter. Acta Biotheoretica 59(3-4):213-229. https://dx.doi.org/10.1007/s10441-011-9138-9

Gregorius H-R (1991) On the concept of effective number. Theoretical Population Biology 40(2):269-283.

Gregorius H-R, B Degen and A König (2007) Problems in the analysis of genetic differentiation among populations – a case study in Quercus robur. Silvae Genetica 56(3-4):190-199. https://www.thuenen.de/media/institute/fg/PDF/Silvae_Genetica/2007/Vol._56_Heft_3-4/56_3-4_190.pdf

Gregorius H-R, EM Gillet and M Ziehe (2003) Measuring differences of trait distributions between populations. Biometrical Journal 45(8):959-973. https://dx.doi.org/10.1002/bimj.200390063

Gregorius H-R and JH Roberds (1986) Measurement of genetical differentiation among subpopulations. Theoretical and Applied Genetics 71(6):826-834. https://dx.doi.org/10.1007/BF00276425

Hedrick PW (2005) A standardized genetic differentiation measure. Evolution 59(8):1633-1638. https://dx.doi.org/10.1111/j.0014-3820.2005.tb01814.x

Hufford KM and JLH Hamrick (2003) Viability selection at three early life stages of the tropical tree Platypodium elegans (Fabaceae, Papilionoideae). Evolution 57(3):518-526. https://dx.doi.org/10.1111/j.0014-3820.2003.tb01543.x

Jost L (2008) G$_{ST}$ and its relatives do not measure differentiation. Molecular Ecology 17(18):4015-4026. https://dx.doi.org/10.1111/j.1365-294X.2008.03887.x

Lotterhos KE and MC Whitlock (2014) Evaluation of demographic history and neutral parameterization on the performance of F$_{S}$T outlier tests. Molecular Ecology 23(9):2178-2192. https://dx.doi.org/10.1111/mec.12725

Müller-Starck G (1993) Auswirkungen von Umweltbelastungen auf genetische Strukturen von Waldbeständen am Beispiel der Buche (Fagus sylvatica L.). Schriften aus der Forstlichen Fakultät der Universität Göttingen und der Niedersächsischen Forstlichen Versuchsanstalt, Bd. 112. Frankfurt am Main, Germany: J.D. Sauerländer's Verlag, 163 p.

Nei M (1973) Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences U.S.A. 70(12):3321-3323. http://www.pnas.org/content/70/12/3321.abstract

Nosil P, DJ Funk and D Ortiz-Barrientos (2009) Divergent selection and heterogeneous genomic divergence. Molecular Ecology 18(3):375-402. https://dx.doi.org/10.1111/j.1365-294X.2008.03946.x

Simpson EH (1949) Measurement of diversity. Nature 163(30 April 1949):688. https://dx.doi.org/10.1038/163688a0

Wright S (1978) Evolution and the Genetics of Populations, Vol. 4: Variability Within and Among Natural Populations. Chicago, The University of Chicago Press, 590 p, ISBN 9780226910529