

Impact of Amplified Fragment Length Polymorphism Size Homoplasmy on the Estimation of Population Genetic Diversity and the Detection of Selective Loci

Armando Caballero,¹ Humberto Quesada and Emilio Rolán-Alvarez

Departamento de Bioquímica, Genética, e Inmunología, Facultad de Biología, Universidade de Vigo, 36310 Vigo, Spain

Manuscript received October 10, 2007
Accepted for publication February 22, 2008

Computer simulations have allowed surveying a larger number of markers by modeling the distribution of fragment lengths and generating probabilities of comigration and homoplasy. The size distribution of AFLP fragments is strongly asymmetrical, with a much higher proportion of smaller than larger fragments, particularly so for genomes with low GC content (INNAN *et al.* 1999; VEKEMANS *et al.* 2002; KOOPMAN and GORT 2004; GORT *et al.* 2006). Thus, the possible impact of homoplasy is larger for small- than for large-size fragments. For example, VEKEMANS *et al.* (2002) analyzed AFLP fragments from samples of one population of *Phaseolus lunatus* and two populations of *Lolium perenne* and divided the observed fragments into four size categories (75–124, 125–199, 200–299, and 300–450 bp). From this experimental data and from simulated data, they found a highly significant negative correlation between fragment size and the frequency of the marker allele associated with the fragments, which suggests that there is an upward bias in the allelic frequency for the small-size fragments. From the empirical data they also showed that the estimates of nucleotide diversity within populations of each species and those of nucleotide divergence between populations of *L. perenne* increased with increasing fragment sizes, suggesting that homoplasy produces an underestimation of both within- and between-population diversity estimated from AFLP fragments.

KOOPMAN and GORT (2004) found that size homoplasy determines a lack of concordance between the empirical AFLP fragment-length distributions and the predicted distributions generated *in silico* for the *Arabidopsis thaliana* genome. A general observation is that the larger the number of bands detected per primer combination of selective nucleotides, the larger the possible magnitude of homoplasy (VEKEMANS *et al.* 2002; KOOPMAN and GORT 2004; GORT *et al.* 2006). This is in agreement with the observations of ALTHOFF *et al.* (2007), who used a computer program that mimics the AFLP procedure to examine the homology of fragments in whole genomes from eight species representing a range of genome sizes. They found that, for a given primer combination, both the number of fragments and the degree of homoplasy increased with genome size. The average number of homoplasious fragments was 11% for small genomes (<400 Mb), but increased to 41.5% for large genomes (>2 Gb).

The empirical and simulation studies described above indicate that homoplasy is a real component of AFLP data sets. This suggests that homoplasy may produce biases in the estimation of genetic diversity from AFLP markers, but the magnitude and direction of the bias remains to be investigated in detail. In addition, the use of AFLPs to identify selective loci in genome scans is becoming an active area of research in evolutionary studies (WILDING *et al.* 2001; STORZ and NACHMAN 2003; CAMPBELL and BERNATCHEZ 2004; SCOTTI-SAINTAGNE

et al. 2004; STORZ *et al.* 2004; STORZ and DUBACH 2004; ACHERÉ *et al.* 2005; BEAUMONT 2005; NIELSEN 2005; STORZ 2005; BONIN *et al.* 2006; JUMP *et al.* 2006; MEALOR and HILD 2006; MURRAY and HARE 2006; MILLER *et al.* 2007; PAPA *et al.* 2007; NOSIL *et al.* 2008; SMITH *et al.* 2008). Genomewide scans of selective loci generally use information on heterozygosity and genetic differentiation (BEAUMONT 2005; STORZ 2005), and the putative biases in these parameters due to homoplasy could produce some misinterpretations. The purpose of this work was to carry out a comprehensive theoretical analysis on the impact of fragment-size homoplasy upon the estimation of genetic diversity within and between populations. We also investigated the performance of a currently used method (DFDIST software) to detect selective loci from the comparison between genetic differentiation and heterozygosity of dominant molecular markers, as well as the impact of AFLP homoplasy on its effectiveness.

ANALYTICAL AND SIMULATION PROCEDURE

An analytical procedure was used to simulate either neutral or selected AFLP fragments in a subdivided population. Genetic-diversity parameters of the population were calculated for all simulated AFLP fragments and for those that would actually be observed in an experimental setting, where fragments of the same size comigrate in the same electrophoretic band (homoplasy). This allowed us to assess the bias imposed by homoplasy on the estimation of the genetic-diversity parameters in the population. An independent method based on coalescent simulations of DNA sequences was used to check the analytical procedure. In the second part of the study, a program used to detect selective loci, based on estimates of population genetic diversity, was used to check the impact of homoplasy on the detection of selective AFLP markers. Each of these procedures is detailed below.

Allele frequency distributions for neutral loci: A classical island model at equilibrium between migration and drift was used to obtain the frequency of neutral markers in a subdivided population. For biallelic loci the distribution of allelic frequencies $\phi(p)$ followed a beta distribution,

$$\phi(p) = \frac{\Gamma[\theta]}{\Gamma[\theta\bar{p}]\Gamma[\theta(1-\bar{p})]} p^{\theta\bar{p}-1} (1-p)^{\theta(1-\bar{p})-1} \quad (1)$$

(WRIGHT 1937), where $\theta = 4Nm[n/(n-1)]$, n is the number of subpopulations, N is the subpopulation size, m is the migration rate, \bar{p} is the average allele frequency in the whole population, and $\Gamma[\cdot]$ is the gamma function. Allelic frequencies sampled from the theoretical distribution (Equation 1) were assigned to each subpopulation for a number (N_L) of loci. The number of subpopulations was assumed to be either $n = 2$ or 100,

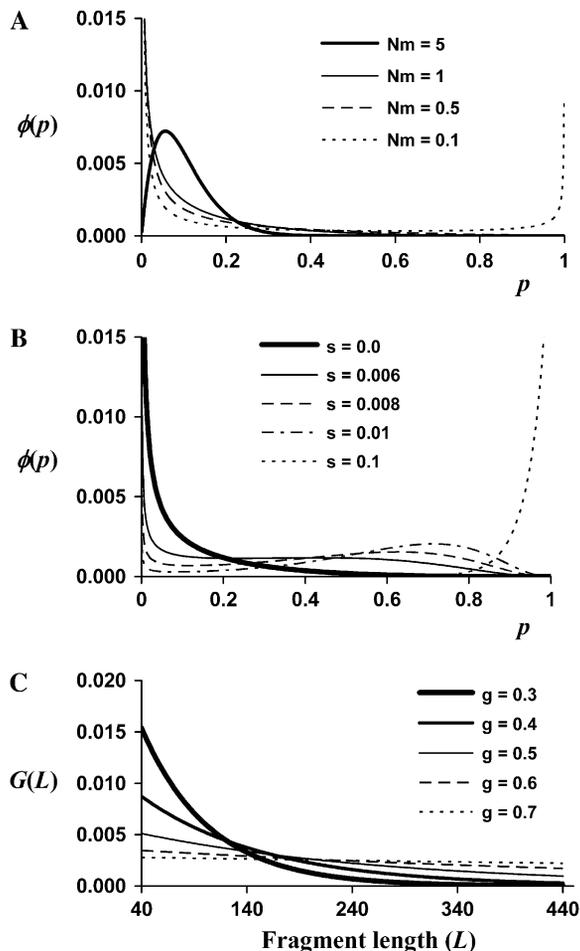


FIGURE 1.—(A) Distribution, $\phi(p)$, of allelic frequencies (p), for a neutral allele in a subdivided population following an island model obtained from the beta distribution (Equation 1; assuming $n = \infty$), for different values of Nm , where N is the subpopulation size and m is the per-generation migration rate. (B) Distribution, $\phi(p)$, of allelic frequencies (p), for a selected allele in a subdivided population obtained from the transition matrix approach, for different values of the selection coefficient of the favorable allele (s), and assuming a number of migrants per generation of $Nm = 1$ from other ($n = \infty$) subpopulations where the average allelic frequency is $\bar{p} = 0.1$. (C) Distribution, $G(L)$, of the number of fragment lengths (L) cut by restriction enzymes *EcoRI* and *MseI*, obtained by Equation 2 of INNAN *et al.* (1999), assuming different values of GC content (g).

and the subpopulation size $N = 500$, although other subpopulation sizes (50, 1000, 5000, and 10,000) were also considered but gave similar results and are not shown. Appropriate values of θ were found so as to produce the desired levels of genetic differentiation (F_{ST} ; WRIGHT 1951) for each scenario. An illustration of the distribution of allelic frequencies for different levels of migration is given in Figure 1A.

Assignment to AFLP fragments: AFLP fragment lengths, obtained from the expected distribution derived by INNAN *et al.* (1999), were randomly ascribed to the N_L loci. For typical fragments flanked by an *EcoRI*

adapter on one side, and an *MseI* adapter on the other, the expected frequency of occurrence of fragments of length L is

$$G(L) = \frac{(1-A)A^L}{A^{L_{\min}} - A^{L_{\max} + 1}}, \quad (2)$$

where $A = (1 - a_1)(1 - a_2)$, $a_1 = (g/2)^2((1 - g)/2)^4$, $a_2 = ((1 - g)/2)^4$, L_{\min} and L_{\max} are the minimum and maximum fragment lengths screened, and g is the genomic GC content (INNAN *et al.* 1999; VEKEMANS *et al.* 2002). The distribution obtained from Equation 2 also applies to fragments flanked by an *EcoRI* adapter on each side (INNAN *et al.* 1999). Values of $L_{\min} = 40$ and $L_{\max} = 440$ were assumed in all simulations that correspond to PCR fragments between 72 and 472, because the typical *EcoRI* and *MseI* primers contain a 16-bp-long sequence when excluding the selective nucleotides. Values of g between 0.3 and 0.7 were investigated. Figure 1C shows the distribution of fragment lengths for the different values of g . Note that the lower the GC content assumed, the larger the number of fragments of short length.

True loci vs. experimental fragments: Irrespective of the length of the fragments, all simulated loci were obviously identified individually. Thus, true fragment frequencies and differentiation could be calculated assuming Hardy–Weinberg equilibrium. This group of fragments is called *true* fragments or *loci* hereafter. In an experimental setting, however, PCR fragments (bands) may correspond to different loci with the same fragment length (homoplasmy). Fragments detected as different PCR bands are called *experimental fragments* or *bands* hereafter. The number of experimental bands depends on the number of true loci and the distribution of fragment sizes (Equation 2). For example, if $g = 0.5$, for $N_L = 1000$ true loci there are ~ 345 experimental bands, whereas for $N_L = 200$ true loci there are ~ 151 experimental bands (see also Figure 4 of VEKEMANS *et al.* 2002 for a graph showing the expected number of detectable fragments as a function of the total number of true fragments). The number of true loci considered varied from 60 to 1000.

A comparison between true and experimental fragments allows us to understand the bias due to homoplasmy when estimating allele frequencies, gene diversity, and differentiation. The frequency of the hypothetical allele that explains the incidence of the l th experimental fragment of length L in a given subpopulation was calculated as

$$p_l^* = 1 - \sqrt{\prod_l (1 - p_l)^2}, \quad (3)$$

where p_l is the frequency of the marker allele associated to the true l th locus in the subpopulation, and the product is for all true loci with the same fragment length L . Allelic frequencies, average expected heterozygosity

ties of subpopulations (H_S), expected heterozygosities for the whole population (H_T), and degree of differentiation, $F_{ST} = (H_T - H_S)/H_T$ (NEI 1973), where H_T and H_S are averages over loci, were calculated for both true and experimental fragments. Results for these parameters were generally presented as the ratio between their corresponding values for experimental and true loci, *i.e.*, as the bias incurred by the experimental fragments with respect to the true loci in the estimation of genetic diversity.

Coalescent simulations: To check the previous analytical procedure by a different, independent method, simulations were carried out by the retrospective coalescent approach. The program SIMCOAL 2.1.2 (EXCOFFIER *et al.* 2000; <http://cmpg.unibe.ch/software/simcoal2/>) was used to simulate DNA sequences corresponding to two subpopulations of equal size ($2N = 1000$ haploid individuals) separated for a period of 500 generations and carrying out reciprocal migration at a rate m per generation. Twenty sequences were sampled from each subpopulation and cut in fragments at the motifs GAATTC (enzyme *EcoRI*) and TTAA (enzyme *MsdI*). The number of nucleotides simulated was chosen so as to produce a total number of ~ 1000 fragments (true loci). In particular, 34,000 nucleotides were required (distributed in four chromosomes with 8600 nucleotides each). The default transition rate of 0.333 and no recombination rate within chromosomes were used. To compare this coalescent method and the previous approach for a particular common scenario, a migration rate $m = 0.0008$ between the two subpopulations and a mutation rate per nucleotide $u = 0.00011$ were found to produce an average allelic frequency of true loci of $\bar{p} = 0.1$ and a degree of differentiation between subpopulations of $F_{ST} = 0.1$. Allelic frequencies, average expected heterozygosities, and degree of differentiation were calculated from the samples for both true and experimental fragments. One hundred coalescent simulations and 100 analytical runs were carried out to compare the results of both methods.

Selective loci: Because Equation 1 can provide the distribution of allelic frequencies only in the case of neutral loci, a transition matrix approach was used to obtain allelic distributions for selective loci (see, *e.g.*, FERNÁNDEZ *et al.* 2005). Assume that for a given locus with genotypes AA , Aa , and aa and genotypic frequencies p^2 , $2pq$, and q^2 , the corresponding fitnesses in a given subpopulation of size N are $1 + s$, $1 + s/2$, and 1 , respectively, with $s \geq 0$. Assume also that the locus is not selected in the remaining subpopulations, where its average frequency is \bar{p} , and that every generation a fraction m of genes arrives at the considered subpopulation by migration. The model is also equivalent to an island–continent model where selection occurs exclusively in the island. To generate the distribution of frequencies for the selected gene in the considered subpopulation, taking account of selection, drift, and migration, a

transition matrix \mathbf{A} , of size $(2N + 1) \times (2N + 1)$, was constructed with terms

$$a_{ij} = \binom{2N}{j} (p')^j (1 - p')^{2N-j},$$

giving the probability that the number of copies of the gene is j at a generation, given that it was i in the previous one, where $p = i/2N$ ($0 \leq i \leq 2N$) and

$$p' = \frac{p^2(1 + s) + pq(1 + s/2)}{1 + 2pqs + p^2s} (1 - m) + \bar{p}m$$

is the expected allele frequency after selection and migration. The state of the vector of frequencies at generation t , $\mathbf{v}_t[i]$, with terms from $i = 0$ to $2N$, was obtained from that at generation $t - 1$ as $\mathbf{v}_t[i] = \mathbf{v}_{t-1}[i]\mathbf{A}$. Iterations were run until reaching the asymptotic stage. This was assumed to occur when the squared difference between two consecutive values of the sum of terms for the vector $\mathbf{v}_t[i]$ was $< 10^{-6}$. Selective coefficients were sampled from an exponential distribution, $f(s) = (1/\bar{s}) \exp(-s/\bar{s})$, with mean effect \bar{s} , and were assigned to randomly chosen fragments. Figure 1B illustrates the distribution of frequencies for a range of selective coefficients. The case corresponding to a neutral locus ($s = 0$) produced the same distribution as that obtained from the beta distribution (Equation 1), as expected (*cf.* line for $Nm = 1$ in Figure 1A and line for $s = 0$ in Figure 1B). For increasing values of the selection coefficient the distribution of allelic frequencies is increasingly displaced to the right, but note that for moderate values of s , the expected frequency can take a wide range of values.

Identification of selective fragments by population comparisons: To assess the impact of homoplasy on the identification of selective loci in AFLP genome scans, data obtained with the analytical approach above was analyzed using the DFDIST software (<http://www.rubic.rdg.ac.uk/~mab/stuff/>), a modification for dominant markers of the software developed by BEAUMONT and NICHOLS (1996). Briefly, this program implements a Bayesian method developed by ZHIVOTOVSKY (1999) to estimate allelic frequencies from the proportion of recessive phenotypes in the sample. It also estimates the WEIR and COCKERHAM (1984) F_{ST} between the subgroups defined in the sample. Coalescent simulations are then performed to generate an F_{ST} null sampling distribution from upon the neutral expectations. The simulated F_{ST} data are used to identify those loci that may not fit the neutral drift model given their unusually low or high F_{ST} values. The software DFDIST is the method most frequently used when using AFLPs (see SCOTTI-SAIN-TAGNE *et al.* 2004; ACHERÉ *et al.* 2005; BONIN *et al.* 2006; JUMP *et al.* 2006; MEALOR and HILD 2006; MURRAY and HARE 2006; MILLER *et al.* 2007; PAPA *et al.* 2007; NOSIL *et al.* 2008; SMITH *et al.* 2008).

For the scenarios to be analyzed with the DFDIST program, we considered two subpopulations of size $N =$

TABLE 1
Experimental setting for the analysis with the DFDIST program

N_L per primer ^a	Selective loci (5%) ^b	Bands per primer ^c	% of homoplasious bands ^d	No. of runs (primers)	Total no. of bands (true or experimental) ^e	Expected no. of bands with at least a selective locus	
						True bands	Experimental bands
1000	50	345	74	3	1035	52	150
200	10	151	26	7	1057	53	70
100	5	87	14	12	1044	52	60
60	3	55	9	19	1045	52	57

^aNumber of true loci assumed per simulation replicate (equivalently to per primer combination).

^bNumber of selective loci per primer combination (5% of N_L).

^cNumber of experimental bands observed per primer combination (average over 100 replicates assuming a GC content of $g=0.5$).

^dPercentage of homoplasious experimental bands, *i.e.*, percentage of experimental bands that include two or more true loci.

^eProduct between number of bands per primer combination and number of primer combinations.

500 and a proportion of loci (between 1 and 10%) assumed to be under positive selection in one of them. For example, for the case of $N_L = 1000$ loci, 50 selective loci and 950 neutral ones were assumed for a proportion of selective loci of 5%. Subpopulation sample frequencies for true fragments (p_s) in a random sample of M individuals taken from the subpopulation were obtained from the binomial probability of sampling j copies of a given allele,

$$\varphi_M(j) = \binom{2M}{j} p^j (1-p)^{2M-j},$$

where p is the subpopulation allele frequency for the locus. Absolute frequencies of theoretical true PCR bands (corresponding to true fragments) and no bands in each sample were calculated as $M \times [p_s^2 + 2p_s(1-p_s)]$ and $M \times [(1-p_s)^2]$, respectively. The analogous absolute frequency of experimental PCR bands (or no bands) was obtained with the corresponding experimental allelic frequencies (p_s^* ; Equation 3). Assumed subpopulation sample sizes were $M = 20, 40, 80,$ and 160 individuals.

The parameter conditions used in the DFDIST analyses were as follows:

1. The critical frequency for the most common allele was 0.99 (loci where the most frequent allele had a frequency ≥ 0.99 were excluded).
2. The scale for the ZHIVOTOVSKY (1999) parameters for estimating allele frequencies was 0.25 (it was checked that the estimation of frequencies was accurate with this parameter value).
3. The number of resamplings used to obtain the confidence intervals for outliers was 10,000 (several runs were made with 100,000 resamplings and results did not change).
4. The smoothing proportion used was 0.04.
5. The estimate of average F_{ST} to be used in the DFDIST simulations was obtained in three ways: (i) the average

estimated F_{ST} calculated by the Ddatacal program (one of the programs of the Dfdist package); (ii) a trimmed mean F_{ST} (provided also by the Ddatacal program) obtained excluding 30% of the highest and 30% of the lowest F_{ST} values (this trimmed mean F_{ST} is supposed to be an estimate of the average “neutral” F_{ST} uninfluenced by outlier loci; BONIN *et al.* 2006); and (iii) the median of F_{ST} values from the file data_fst_outfile provided by the Ddatacal program, after removing those loci for which the heterozygosity is < 0.1 . This was the procedure used in the simulations of BEAUMONT and BALDING (2004).

6. Different values of the parameter $4N_e\mu$ were used (from 0.004 to 0.1), finding no differences in the performance of the method within this range. Thus, a value of 0.04 was used throughout.
7. The critical probability for detecting outliers was $\alpha = 5$ or 0.05%.

To make fair comparisons between true and experimental data sets, a given number of runs were pooled in each case to provide the same number of true and experimental loci (bands) for analysis with the program DFDIST (see Table 1). For example, for the case with $N_L = 1000$ true fragments, for which only ~ 345 experimental fragments become available from each run because of fragment-size homoplasmy, three simulation replicates (that would be equivalent to using three different primer combinations in a PCR setting) were pooled. From each replicate, all experimental fragments (~ 345) and 345 randomly taken true fragments were used. Thus, $\sim 345 \times 3 = 1035$ true or experimental loci were analyzed by the DFDIST program and compared. It should be noted that although the proportion of the genome with selective loci is 5%, the 1035 experimental fragments implied 15% of selective loci because each experimental fragment represents on average three true loci (*i.e.*, the 1035 experimental fragments actually correspond to $\sim 1035 \times 3 = 3105$ true loci). An analogous

TABLE 2

Results obtained with the analytical approach (Equations 1–3) and the simulation approach (using the program SIMCOAL; see text for details) for the average number of loci, average allele frequency (\bar{p}), expected heterozygosity within subpopulations (H_S), and genetic differentiation between subpopulations (F_{ST}) for true and experimental fragments

	Analytical approach			Simulation approach (SIMCOAL)		
	True	Experimental	Bias	True	Experimental	Bias
No. loci	1000.0	345.5		1000.5	351.3	
\bar{p}	0.100	0.251	2.503	0.102	0.249	2.445
H_S	0.147	0.286	1.941	0.140	0.266	1.894
F_{ST}	0.100	0.090	0.893	0.100	0.092	0.925

Bias is the ratio between true and experimental estimates for the different parameters. Results are based on 100 replicates for each approach. Standard errors <0.002 for \bar{p} , H_S , and F_{ST} and <0.4 (Analytical) and 6.1 (SIMCOAL) for the number of loci are shown. Parameters: $N_L = 1000$, subpopulation size $N = 500$, number of subpopulations $n = 2$, genome GC content $g = 0.5$, average allelic frequency $\bar{p} = 0.1$, and neutral genetic differentiation $F_{ST} = 0.1$.

procedure was carried out for other numbers of true loci (see Table 1).

Analyses were carried out for different values of F_{ST} and mean selection coefficients ($\bar{s} = 0, 0.001, 0.005, 0.05, \text{ and } 0.5$). The proportion of outliers detected by the method (loci above the 95 or 99.95% confidence limit) and the proportion of true selective loci detected were calculated both for true and experimental loci. Each scenario was replicated 10 times and results were averaged over replicates.

RESULTS

Impact of homoplasmy on neutral gene diversity and differentiation among subpopulations: Table 2 compares the results obtained with the analytical method and the simulation (SIMCOAL) method with the objective of checking the former. The scenario corresponds to a population subdivided into two subpopulations ($n = 2$) of size $N = 500$ each and a migration rate such that the differentiation between subpopulations is $F_{ST} = 0.1$. The average allelic frequency in both cases was set up to be $\bar{p} = 0.1$ for the true loci. The first column for each approach refers to the true loci, the corresponding results for the experimental “loci” are shown in the second column, and the ratio between both (*i.e.*, the amount of bias of the experimental estimates relative to the true values) are in the third column. In this particular case, homoplasmy produces a 2.5-fold inflation in the average allelic frequency (\bar{p}), a 1.9-fold inflation in gene diversity (expected heterozygosity within subpopulations, H_S), and a 10% reduction in the value of genetic differentiation (F_{ST}). All results were very similar for both approaches. In what follows, all results will refer to the analytical method, which was faster and easier to implement and allowed for better control of the parameters.

Figure 2 shows the bias for the parameters estimated with the experimental fragments relative to the true loci

for a range of values of genetic differentiation between subpopulations (F_{ST}) for a scenario of $N_L = 1000$ true loci (continuous line) and 200 loci (broken line). These correspond to ~345 and 151 experimental fragments, *i.e.*, observed PCR bands (see Table 1). For 1000 true fragments the biases in the average allelic frequency and heterozygosity are very substantial and almost invariable

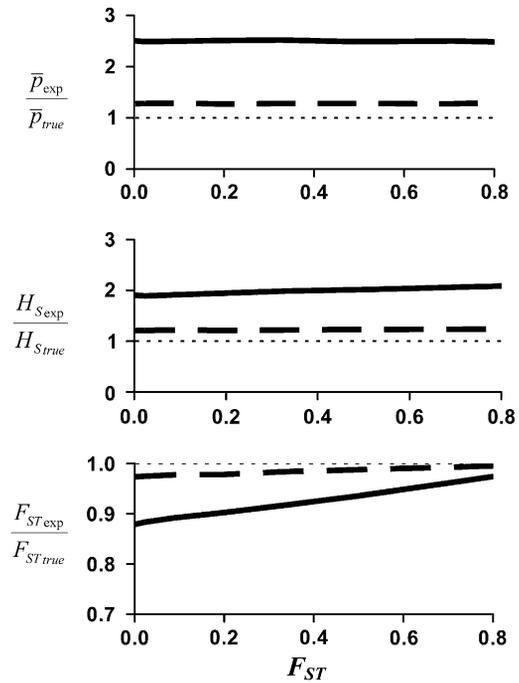


FIGURE 2.—Amount of bias of the estimates of average allelic frequency (\bar{p}), average expected heterozygosity (H_S), and degree of genetic differentiation between subpopulations (F_{ST}) obtained from experimental fragments relative to true loci, for different values of F_{ST} . The continuous line refers to the case with $N_L = 1000$ true loci, and the broken line to $N_L = 200$. The average allelic frequency of true fragments is $\bar{p} = 0.1$. The default parameters are $N = 500$, $n = 100$, $g = 0.5$, and $\bar{p} = 0.1$. All results are the averages of 10 replicates.

for different values of F_{ST} . The bias in the estimated level of differentiation is maximal at low values of F_{ST} and decreases at higher ones. Note, however, that this bias is a relative one, and the absolute bias (absolute difference in F_{ST} between experimental and true fragments) is higher for intermediate values of F_{ST} . All biases are substantially lower with 200 true fragments, as the degree of homoplasy is relatively small in this case.

Figure 3 shows the bias of the experimental parameters relative to the true ones under a range of situations. The first column of graphs shows the biases for different values of the GC content (g). The continuous line refers to $N_L = 1000$ true loci (345 bands) and the broken line to 200 loci (151 bands). The graphs show that the bias is generally larger for small values of g , as expected, but the variation is not very large. The second column of graphs shows results for different average allele frequencies in the whole set of subpopulations (\bar{p}). The lower the true average frequency, the higher is the bias for the esti-

m1.19r6008.(tagi-)ti1.19r7.2(1)-282.4(para007)8(ten..19r7.2(F)-40853that)e.19r60ha007ozerregofdt (sr20e the frequ sy

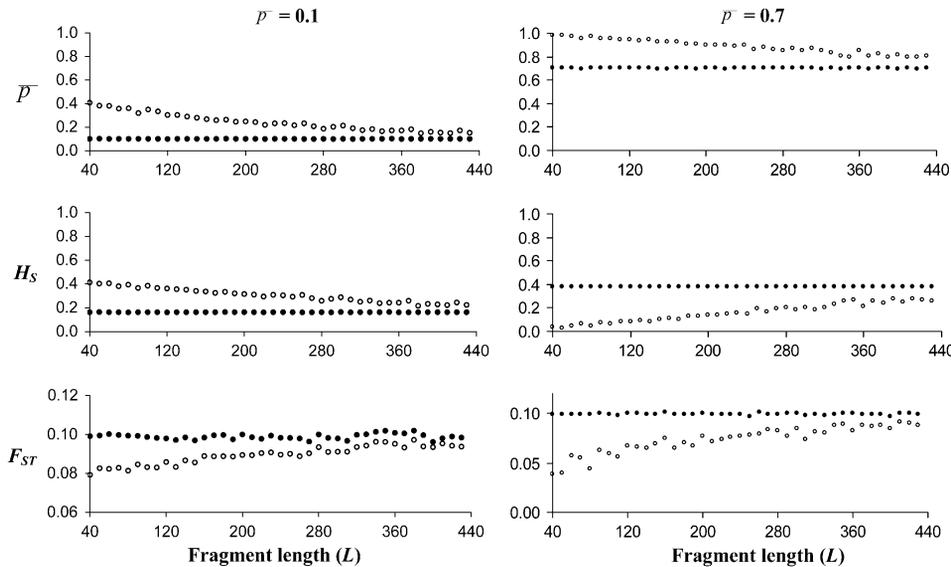


FIGURE 4.—Average allelic frequency (\bar{p}), average expected heterozygosity (H_S), and degree of genetic differentiation between subpopulations (F_{ST}) from experimental fragments (open circles) and true fragments (solid circles), for different classes of fragment lengths. The results correspond to one replicate with the following parameters: $N = 500$, $n = 100$, $g = 0.5$, $F_{ST} = 0.1$, and $\bar{p} = 0.1$ (left column of graphs) or $\bar{p} = 0.7$ (right column of graphs).

erozygosity are overestimated by $\sim 15\%$. Thus, in terms of estimation of neutral genetic differentiation, homoplasy has a substantial impact only when $> \sim 200$ experimental bands per primer are scored. However, in terms of estimation of heterozygosity, a substantial impact may occur even for 50–75 experimental bands scored per primer combination.

Detection of selective loci: *Assessment of the efficiency of the DFDIST software for detecting selective loci:* We first assessed the efficiency of the DFDIST software for detecting selective loci considering only the true loci, *i.e.*, without yet introducing the problem of homoplasy. Figure 5 shows results for a variable proportion of selective loci assumed in the genome (from 0 to 10%), considering an average selection coefficient for the selective genes of $\bar{s} = 0.5$ or 0.005, a critical probability value $\alpha = 5\%$, and an average neutral $F_{ST} = 0.025$. Note that this latter (as before, and thereafter, unless indicated) refers to NEI's (1973) estimate of F_{ST} [the corresponding WEIR and COCKERHAM's (1984) estimate used by the DFDIST software is 0.037]. The left column of graphs in Figure 5 represents the percentage of outliers detected by the method. This is $\sim 4\%$ when the nontrimmed (NT) F_{ST} mean is assumed, $\sim 10\%$ when the trimmed (T) F_{ST} mean is assumed, and $\sim 15\%$ when the median (M) F_{ST} is assumed, and these results are almost independent of the proportion of selective loci simulated and the average selection coefficient. The middle column of graphs in Figure 5 shows the percentage of outliers that are in fact selective loci. This percentage is substantial only for a high proportion of selective loci in the genome, a large \bar{s} and the use of the trimmed or nontrimmed mean F_{ST} . Finally, the right column of graphs in Figure 5 shows the percentage of selective loci that are actually detected by the method. This is very high for selective loci with $\bar{s} = 0.5$ but on the order of $\sim 10\text{--}20\%$ for small \bar{s} , almost irrespective of the proportion of selective loci in the genome.

Figure 6 presents results analogous to those of Figure 5 but using a critical probability value of $\alpha = 0.05\%$ to detect outliers. The percentage of outliers detected gives an approximately right estimate of the true proportion of selective loci in the genome only for scenarios with a large \bar{s} , a large proportion of selective loci ($\geq 3\%$), and using a trimmed mean or median F_{ST} (Figure 6, G and M). Otherwise, the prediction is rather deficient. The percentage of selective loci present in outliers is very high for $\bar{s} = 0.5$ in most cases, except for a large proportion of selective loci (10%) and a nontrimmed mean F_{ST} (Figure 6B; for which no outliers are in fact detected) and when the proportion of selective loci is small (1%) with a trimmed or median F_{ST} (Figure 6, H and N). The percentage of selective loci in outliers is rather small for $\bar{s} = 0.005$, with a better performance of the mean F_{ST} (Figure 6, E and K) than the median F_{ST} (Figure 6Q). The percentage of selective loci detected is large only for high \bar{s} using the trimmed mean or the median F_{ST} (Figure 6, I and O).

Figure 7 extends the above results assuming a trimmed mean F_{ST} and 5% of selective loci in the genome for a wider range of average selection coefficients, a different value of neutral F_{ST} (0.1), and different sample sizes (M). The percentage of outliers detected with a critical probability of $\alpha = 5\%$ is again $\sim 10\%$ for $F_{ST} = 0.025$ (Figure 7A), irrespective of the average selection coefficient, and this value is reduced to $\sim 7\%$ for $F_{ST} = 0.1$ (Figure 7D). Runs corresponding to a higher value of $F_{ST} = 0.2$ actually produce even lower proportions of outliers (data not shown). For a critical probability of $\alpha = 0.05$ (Figure 7, G and J) the correct proportion of selective loci (5%) is predicted only for a large average selection coefficient ($\bar{s} = 0.5$) and a small level of genetic differentiation ($F_{ST} = 0.025$) but not for smaller values of \bar{s} or a larger value of $F_{ST} = 0.1$. The percentage of selective loci found in the outliers is $> 80\%$ only for large

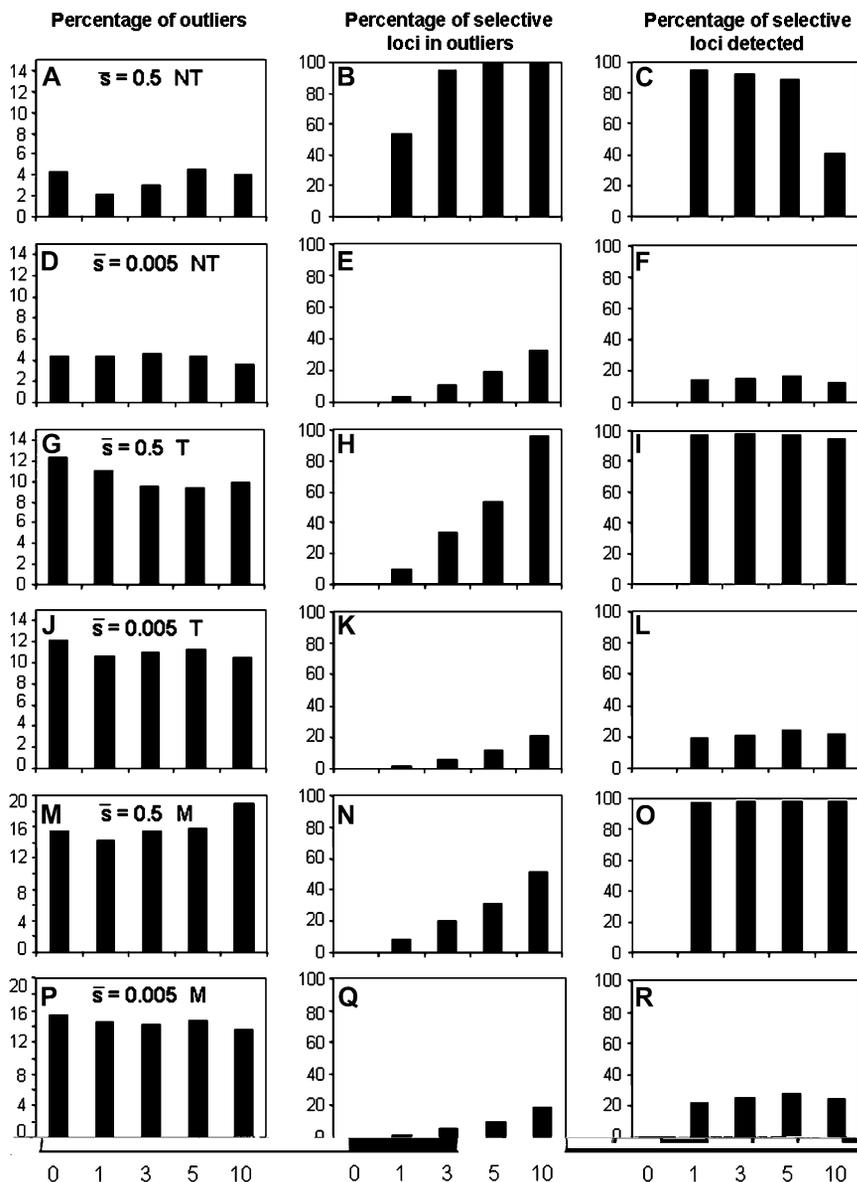


FIGURE 5.—Percentage of outliers (A, D, G, J, M, and P), percentage of outliers that correspond to truly selective loci (B, E, H, K, N, and Q), and percentage of selective loci that are detected as outliers (C, F, I, L, O, and R) by the program DFDIST after analyzing ~ 1000 (true) fragments, for a variable proportion of selective loci in the genome with average effect \bar{s} . The critical probability threshold value for detecting outliers is $\alpha = 5\%$. The average estimate of the F_{ST} used for the DFDIST simulations is the nontrimmed mean (NT), a trimmed mean (T), and the median (M). Default parameters (see text): $N = 500$, $N_L = 1000$, $n = 2$, $g = 0.5$, $\bar{p} = 0.1$, $M = 40$, and $F_{ST} = 0.025$. All results are the average of 10 replicates.

values of \bar{s} (0.05–0.5) and a critical probability of $\alpha = 0.05\%$ (Figure 7, H and K). The percentage of selective loci detected generally increases with the average selection coefficient, as would be expected. Finally, increasing the sample size does not have any impact on the proportion of outliers detected or the efficiency of detection (Figure 7, M–O) beyond a sample size of $M = 40$.

Impact of homoplasy on the detection of selective loci: Because the levels of heterozygosity and F_{ST} are affected by homoplasy (Figures 2–4), and these are the parameters used for the detection of selective loci from molecular data, we investigated the impact of homoplasy on the detection of selective loci by comparing the analyses of true and experimental fragments. Table 3 gives average parameters from true loci and estimated parameters from experimental fragments in a case where 5% of the true loci (*i.e.*, 50 of 1000) have an

average selection coefficient of $\bar{s} = 0.005$. The first part of the table shows parameters including all loci (50 selective and 950 neutral loci). Because of the presence of selective loci, the average allelic frequency (\bar{p}), average heterozygosity (H_S), and average F_{ST} are increased with respect to the purely neutral case (*cf.* Tables 1 and 3). The biases incurred by the experimental fragments are slightly smaller (for \bar{p} and H_S) or larger (for F_{ST}) than for the neutral case (Table 1). The right side of Table 2 shows the biases exclusively for the selected loci. The highest bias from experimental loci occurs as a severe underestimation of F_{ST} , which comes down to about one-half of the true parameter value.

To assess the overall impact of homoplasy, Figure 8 presents results analogous to those of Figures 5–7, but comparing true and experimental fragments. The solid bars refer to the analysis of a set of ~ 1000 fragments that are true loci, whereas the open bars refer to the cor-

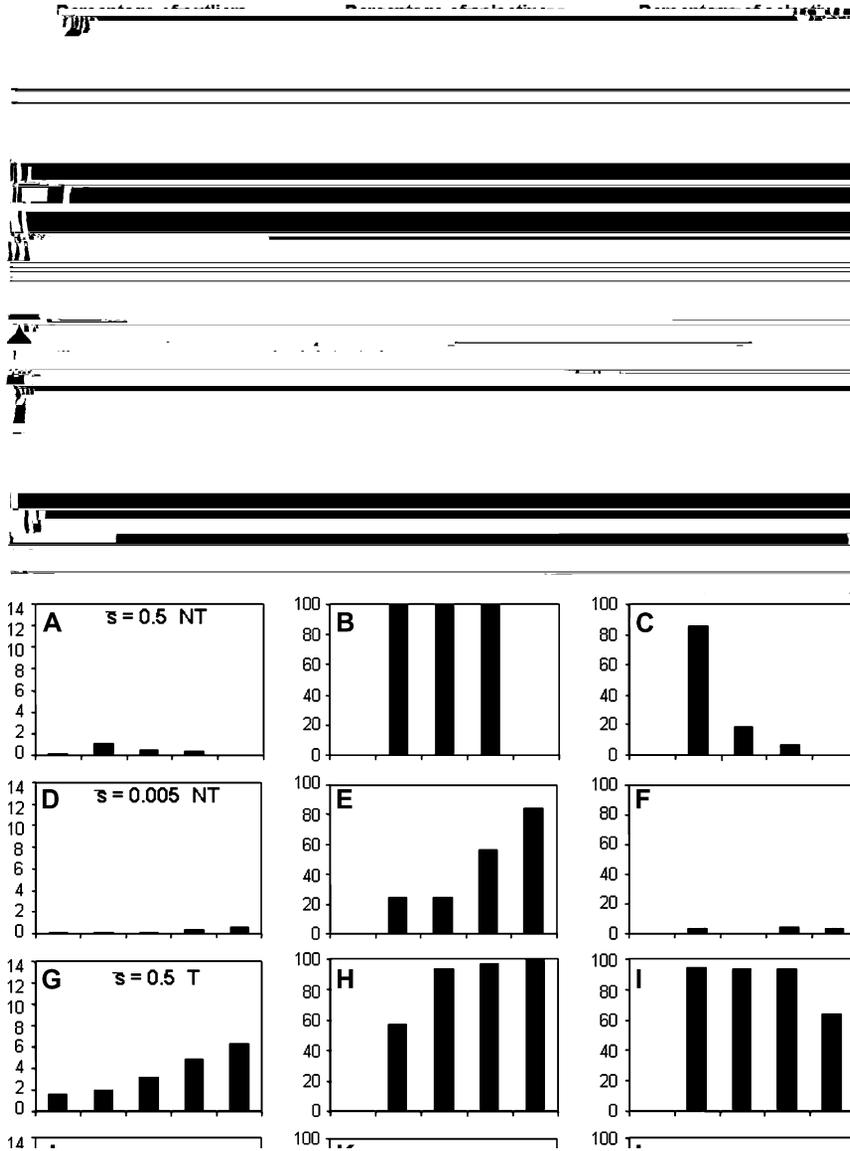


FIGURE 6.—Percentage of outliers (A, D, G, J, M, and P), percentage of outliers that correspond to truly selective loci (B, E, H, K, N, and Q), and percentage of selective loci that are detected as outliers (C, F, I, L, O, and R) by the program DFDIST after analyzing ~ 1000 (true) fragments, for a variable proportion of selective loci in the genome with average effect \bar{s} . The critical probability threshold value for detecting outliers is $\alpha = 0.05\%$. All other parameters and definitions are as in Figure 5.

responding analysis of ~ 1000 fragments that are experimental fragments, hypothetically obtained with a variable number of primer combinations. Thus, for example, the first open bar in each graph (55 bands per primer for $g = 0.5$ or 50 for $g = 0.35$) would refer to a case where there are 19 primers producing 55 (or 50) experimental bands each (see Table 1). The second bar (87 or 76 bands per primer depending on g) corresponds to a case where there are 12 primers producing 87 (or 76) experimental bands each, and so on. Higher levels of homoplasy, corresponding to larger numbers of experimental bands, imply a slight overestimation of the percentage of outliers (Figure 8, A, D, and G) and an underestimation of the percentage of selective loci that are detected (Figure 8, C, F, and I). Homoplasy has no appreciable impact on the percentage of outliers carrying true selective loci (Figure 8, B, E, and H). The overestimation of the percentage of outliers detected can be understood from the information presented in Table 1.

Note that a larger number of experimental bands implies a progressively larger number of true loci (N_L). Thus, for example, for $N_L = 1000$ (implying 345 experimental bands for $g = 0.5$ or 252 for $g = 0.35$), the 1000 experimental bands analyzed correspond to ~ 3000 true loci, from which ~ 150 are selected (5%), as in this case each band includes an average of three true loci of the same fragment length. Thus, $\sim 15\%$ of the experimental fragments could include at least one selective true fragment. This explains why the percentage of outliers is larger for the experimental fragments than for the true loci.

DISCUSSION

AFLP homoplasy and gene diversity and differentiation among subpopulations: Our results demonstrate that the biases due to fragment homoplasy can be rather

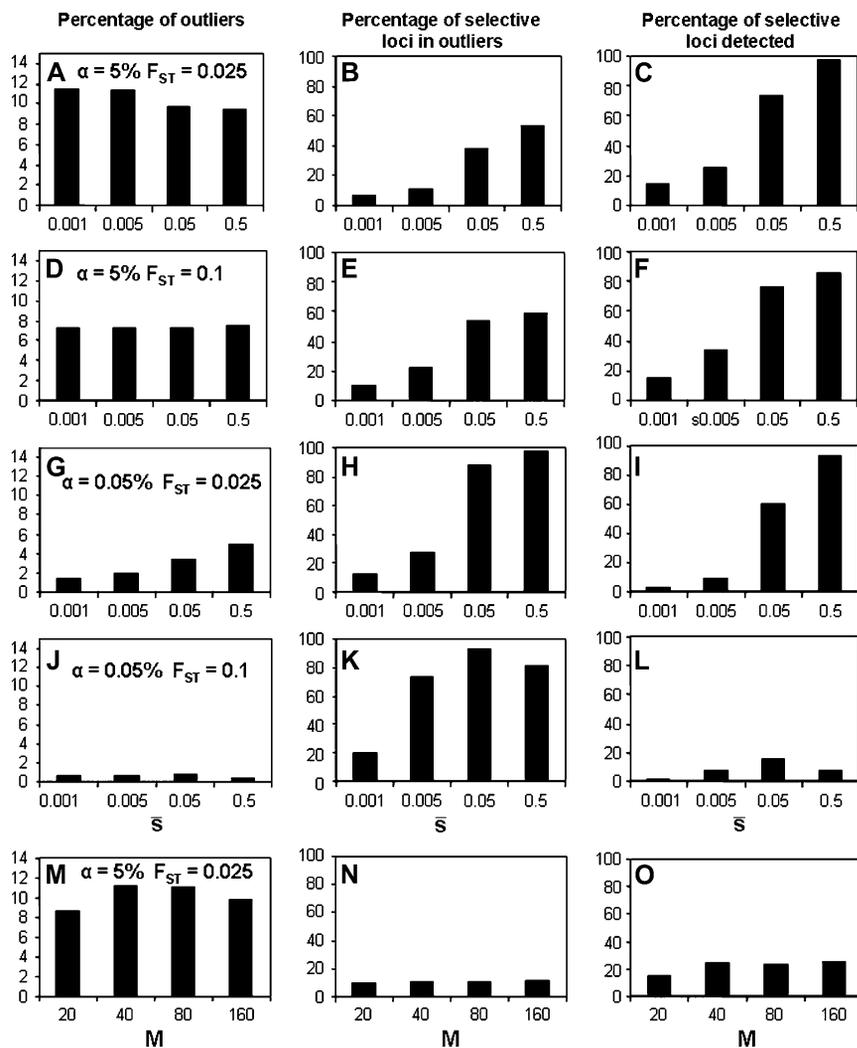


FIGURE 7.—Percentage of outliers (A, D, G, J, and M), percentage of outliers that correspond to truly selective loci (B, E, H, K, and N), and percentage of selective loci that are detected as outliers (C, F, I, L, and O) by the program DFDIST after analyzing ~ 1000 (true) fragments, for a variable average effect of selective loci (\bar{s}) or sample size (M). The critical probability threshold value for detecting outliers is α . F_{ST} , degree of neutral differentiation. A trimmed mean average estimate of F_{ST} is used for the DFDIST simulations. Default parameters (see text): $N = 500$, $N_L = 1000$, $n = 2$, $g = 0.5$, $\bar{p} = 0.1$, and $M = 40$. All results are the average of 10 replicates.

high for the estimates of \bar{p} and H_S , whereas generally modest for the estimates of F_{ST} (Figures 2 and 3). These biases depend not only on technical aspects that can be modified during the assay (number of loci scored per primer and fragment size), but also on factors usually unknown *a priori* in nonmodel organisms, such as the GC genomic composition or the demographic history and structure of populations under study.

Our simulations indicate that the bias associated with size homoplasies rapidly increases with the number of true loci. Vos *et al.* (1995) recommended that, to avoid size homoplasies, the number of fragments or bands in an AFLP profile must be between 50 and 100, whereas GORT *et al.* (2006) suggested that even band numbers as low as 20 are not a guarantee of the absence of band homoplasies. These results are in line with an empirical study on the sugar beet by HANSEN *et al.* (1999), who scored 456 bands in 16 AFLP lanes, giving an average of 28.5 bands per line. They reported that 13.2% of the bands were likely to be homoplasious. In our study, the comparison between the biases associated with the estimation of population parameters with a number of

true loci ranging from 60 to 1000 indicates that ~ 100 experimental fragments per primer combination may represent an upper limit to fragment homology among bands. From a literature review of 90 AFLP surveys on diverse animal, plant, fungi, and bacteria species (in the period 2004–2006), we found that the average number of fragments detected per primer combination ranges from 4 (ENJALBERT *et al.* 2005) to 453 (KALITA and MALEK 2006). If we take the median value of 58.8 as the more common number of detected fragments per primer combination, biases associated with the estimation of population parameters such as heterozygosity, F_{ST} and allele frequency are almost negligible in most studies. However, 18 (20%) studies showed an average number of fragments per primer combination >100 and are expected to have a high risk of biases due to size homoplasies (*e.g.*, LAITUNG *et al.* 2004; SNÄLL *et al.* 2004; ALBACH *et al.* 2006; FANG *et al.* 2006; KALITA and MALEK 2006; KIDD *et al.* 2006; TROELL *et al.* 2006). Furthermore, as shown above, the overestimation of the expected heterozygosity can be substantial (15–25%) even for as low as 50–75 fragments per primer combination.

TABLE 3

Average number of loci, average allele frequency (\bar{p}), expected heterozygosity within subpopulations (H_S), and genetic differentiation between subpopulations (F_{ST}) for true and experimental fragments

	All loci			Selected loci		
	True	Experimental	Bias	True	Experimental	Bias
No. loci	1000.0	346.2		50.0	46.3	
\bar{p}	0.211	0.454	2.148	0.405	0.657	1.620
H_S	0.255	0.338	1.325	0.224	0.273	1.218
F_{ST}	0.134	0.110	0.824	0.434	0.228	0.525

Five percent of the loci (50) are selective with a selection coefficient sampled from an exponential distribution with average $\bar{s} = 0.005$; bias is the true/experimental ratio for the different parameter estimates. Results are based on 10 replicates. SEs < 0.011 for \bar{p} , H_S , and F_{ST} and < 1.6 for N_L are shown. The default parameters are $N_L = 1000$, $N = 500$, $n = 2$, $g = 0.5$, $\bar{p} = 0.1$, and $F_{ST} = 0.1$ (see Table 2).

VEKEMANS *et al.* (2002) demonstrated that smaller-sized fragments are more likely to be homoplasious. Indeed, it is possible that the high degree of monomorphism often observed for small (< 125 bp) fragments could be due to elevated levels of comigration of nonhomologous fragments (VEKEMANS *et al.* 2002; BUSSEL *et al.* 2005). Consistent with this, our results show that the bias due to size homoplasy increases in the small-size classes (Figure 4), although there is no clear cutoff point. Thus, shorter fragments have lower quality data. Inclusion of fragment-length information in the estimation of heterozygosity, \bar{p} , and F_{ST} would likely result in more reliable estimates. Indeed, this approach has been efficiently used to obtain weighted similarity coefficients between unrelated genotypes (KOOPMAN and GORT 2004).

An additional difficulty is that the relationship between fragment size and homoplasy bias is highly dependent on the average allele frequency, leading to either overestimates or underestimates of heterozygosity.

For example, VEKEMANS *et al.* (2002) deduced that, for small-size fragments, homoplasy would produce underestimation of both within- and between-population genetic diversity. This is in agreement with the results observed in Figure 4 with high $\bar{p} = 0.7$, where it can be seen that homoplasy produces underestimations of both H_S and F_{ST} . In fact, the average frequency of the allele producing a band in the empirical data analyzed by VEKEMANS *et al.* (2002) was 0.714 and 0.444 for the two species studied (see Table 2 in their article). However, if the average frequency had been much smaller, heterozygosity would have been overestimated instead of underestimated (as shown in the scenario with $\bar{p} = 0.1$ in Figure 4).

Our analysis demonstrates a moderate influence of the genome GC content on the homoplasy bias. The enzymes employed in this study are a frequent cutter (*MseI*) and a rare cutter (*EcoRI*) commonly used in AFLP surveys. Because *MseI* cuts are more frequent than *EcoRI* cuts, the average AFLP fragment size will be determined

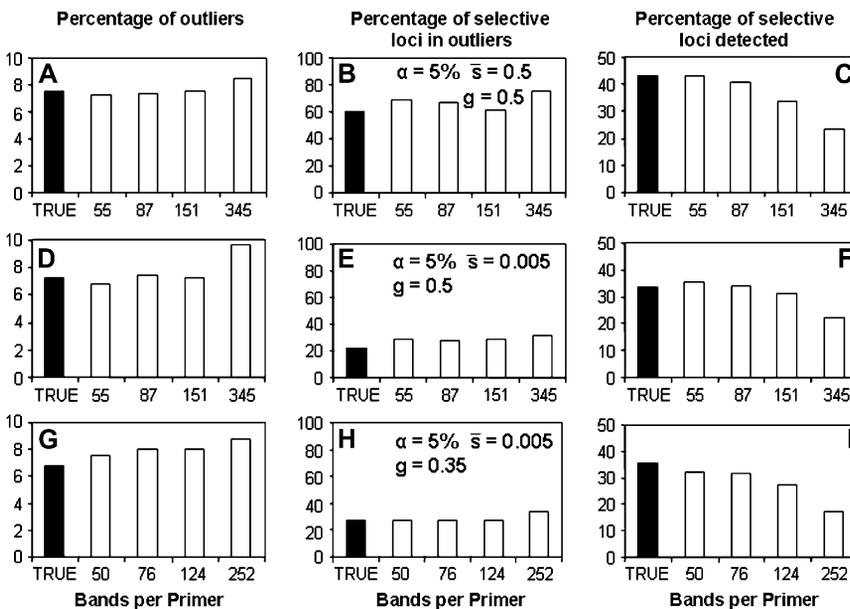


FIGURE 8.—Percentage of outliers (A, D, and G), percentage of outliers that correspond to truly selective loci (B, E, and H), and percentage of selective loci that are detected as outliers (C, F, and I) by the program DFDIST after analyzing ~ 1000 fragments, against number of bands per primer combination. Solid bars correspond to true loci whereas open bars refer to experimental (observed) fragments, which correspond to 19 primer combinations (55 or 50 bands per primer for $g = 0.5$ or $g = 0.35$, respectively), 12 primer combinations (87 or 76 bands), 7 primer combinations (151 or 124 bands), and 3 primer combinations (345 or 252 bands; see Table 1). \bar{s} , average selection coefficient. g , genome GC content. Other default parameters (see text): $N = 500$, $n = 2$, $\bar{p} = 0.1$, $M = 40$, and $F_{ST} = 0.1$. All results are the averages of 10 replicates.

mainly by the frequency of *MseI* cuts. The restriction motif of *MseI* contains no G + C nucleotides, thus increasing the contribution of the small-sized class to the proportion of nonhomologous comigrating bands at low GC contents. This effect has also been observed with experimental data and probably results from the isochore structure of the genome in combination with the restriction motif recognized by the enzymes (*i.e.*, *MseI* will preferably cut in A + T-rich isochores; KOOPMAN and GORT 2004). VEKEMANS *et al.* (2002) found that the substitution of the commonly used *MseI* enzyme by a restriction enzyme recognizing a sequence motif with a balanced GC content would make the AFLP size distribution insensitive to variation in GC content. Hence, the effect of GC content on the homoplasmy bias will depend on the choice of the restriction enzyme and may be avoided by using enzymes containing a balanced GC content for the shortest restriction motif, such as, for example, *TaqI* (VEKEMANS *et al.* 2002). However, the impact of homoplasmy cannot be removed altogether even if a set of enzymes is used that guarantees a flat distribution of fragment lengths. This can be seen in Figure 3. When $g = 0.7$, the distribution of fragment lengths is almost uniform for the *EcoRI*–*MseI* tandem (see Figure 1C) but, even in this case, it is predicted to overestimate \bar{p} and H_S , and underestimate F_{ST} (see Figure 3; left graphs).

The model investigated assumes that the level of homoplasmy depends only on the number of fragments evaluated (see Table 1) and, therefore, is independent of the level of differentiation between populations or the average allele frequency of the markers. It is however possible that increasing differentiation is associated with increasing homoplasmy, as has been repeatedly observed empirically (O'HANLON and PEAKALL 2000; EL-RABEY *et al.* 2002; MECHANDA *et al.* 2004). O'HANLON and PEAKALL (2000) showed a positive correlation of the proportion of nonhomologous bands with species divergence and a negative relationship between the number of comigrating fragments and DNA sequence divergence. For distantly related species, many of the fragments tend to be unique to a species rather than shared among species, and a portion of these would be homoplasious. Thus, as the number of comigrating AFLP bands decreased, the chance that they were non-homologous increased. If this can be generalized, the impact of homoplasmy on the bias of population genetic diversity among distantly related species could be of much higher magnitude than that among conspecific populations.

To avoid the effects of homoplasmy on the estimates of neutral variation, VEKEMANS *et al.* (2002) suggested (i) avoiding fragments in the small-size classes, (ii) using fragments of known map position that should have documented Mendelian inheritance when possible, and (iii) increasing the number of selective nucleotides in the primer sequences to reduce the number of frag-

ments and thus the probability of nonhomology. However, although the exclusion of fragments of small size (*i.e.*, <125 bp) would help to reduce levels of homoplasmy by approximately one-third (GORT *et al.* 2006), there could still be fragments of larger size that are homoplasious. Alternatively, mapping studies are technically demanding, time-consuming, and out of the scope of many AFLP surveys (MEUDT and CLARKE 2007). The reduction of the number of scored bands per AFLP profile is also in conflict with the consideration that very small numbers of bands are undesirable in the sense that they reduce the information content of the AFLP profile. Therefore, even for a low number of bands per lane, a considerable percentage of comigrating bands are still nonhomologous (VEKEMANS *et al.* 2002; GORT *et al.* 2006).

Detection of selective loci from gene diversity and differentiation: On the basis of the classical LEWONTIN and KRAKAUER (1973) multilocus neutrality test, several related approaches have recently been proposed for the detection of selective loci in genome scans (BOWCOCK *et al.* 1991; BEAUMONT and NICHOLS 1996; VITALIS *et al.* 2001; SCHLÖTTERER 2002; BEAUMONT and BALDING 2004). One of the most popular methods, and for which software to analyze dominant data is available (DFDIST), is that from BEAUMONT and NICHOLS (1996). BEAUMONT and BALDING (2004) investigated, by computer simulations, the performance of the codominant version of the method (FDIST program) in comparison with a hierarchical-Bayesian approach, in a structured population scenario for a set of biallelic codominant loci predominantly neutral but some of them subject to directional or balancing (heterozygote advantage) selection. They concluded that both methods could identify loci subject to directional selection when the selection coefficient was at least five times the migration rate, but that neither method was able to identify loci subject to balancing selection. We extended the analysis of BEAUMONT and BALDING (2004) by evaluating the performance of the DFDIST software, considering dominant markers, a distribution of variable selective effects, and a scenario of two subpopulations with different levels of neutral genetic differentiation, where a variable proportion of the loci (that was set up from 1 to 10% of the genome) are subject to directional selection in one subpopulation. We assessed a number of alternatives used in different empirical or theoretical studies:

1. A high critical probability threshold level of 5% *vs.* a much lower value of 0.05%: Most authors have considered critical values of 5% (SCOTTI-SAINTAGNE *et al.* 2004; ACHERÉ *et al.* 2005; BONIN *et al.* 2006; MEALOR and HILD 2006; MILLER *et al.* 2007; PAPA *et al.* 2007; NOSIL *et al.* 2008; SMITH *et al.* 2008) or 1% (JUMP *et al.* 2006; MURRAY and HARE 2006; PAPA *et al.* 2007; NOSIL *et al.* 2008).
2. The mean F_{ST} used to build the null model neutral distribution was a trimmed mean (used by BONIN

et al. 2006; MILLER *et al.* 2007; NOSIL *et al.* 2008; SMITH *et al.* 2008), the raw (nontrimmed) mean (used by SCOTTI-SAINTAGNE *et al.* 2004; MEALOR and HILD 2006), or the median F_{ST} (used in simulations by BEAUMONT and BALDING 2004).

We found that the potentiality of the method to estimate the proportion of the genome subject to selection through the proportion of outliers detected has some reliability only for the case where (i) the selection coefficients are very large ($\bar{s}=0.5$), (ii) the trimmed mean or median F_{ST} are used, (iii) the critical probability level is low ($\alpha = 0.05\%$), and (iv) the level of differentiation is also low ($F_{ST} = 0.025$) (Figures 5–7). In all other situations, such as for $\alpha = 5\%$ (Figure 5), nontrimmed mean or low \bar{s} (Figure 6), or a higher level of differentiation ($F_{ST} = 0.1$; Figure 7), the method does not provide a right inference of the proportion of selective loci in the genome. Even in the most favorable scenarios above, the percentage of outliers does not always predict the true proportion of selective loci in the genome. For example, using the median F_{ST} and $\alpha = 0.05\%$, for $\bar{s} = 0.5$ (Figure 6M), the proportion of outliers is $>3\%$ whereas the true proportion of selective loci is 1%. In accordance with our results, BEAUMONT and BALDING (2004) found a proportion of outliers (as we deduced from their Table 3) close to the true proportion of selective loci in the genome (8% in their case) when considering a high constant coefficient of selection $s = 0.2$. However, the percentage of outliers detected was $\ll 8\%$ for selective loci of small effect ($s = 0.02\text{--}0.05$).

It is highly remarkable that when a high critical probability threshold ($\alpha = 5\%$) is applied, the proportion of outliers detected is basically independent of the proportion of selective loci in the genome (Figure 5) as well as of the value of \bar{s} (Figure 7), which indicates that for relatively low values of \bar{s} or low proportions of selective loci, the majority of outliers are false positives. This produces an obvious sense of lack of reliability in the method. In theory, the trimmed mean F_{ST} or the median F_{ST} should correct for the presence of selective loci and, therefore, they should remain invariable for increasing values of \bar{s} or the proportion of selective loci in the genome. However, this is not the case, and both parameters increase correspondingly, although to a lower extent than the uncorrected mean F_{ST} . For example, for $F_{ST} = 0.025$ and 5% of selective loci in the genome, the trimmed mean WEIR and COCKERHAM'S (1984) F_{ST} is 0.016, 0.018, and 0.038, and the median F_{ST} is 0.011, 0.013, and 0.016 for $\bar{s} = 0, 0.005$ and 0.5, respectively. This increase, inherent to any correction method, may be the cause of an approximately constant proportion of outliers irrespective of the number and effects of selective loci.

In most studies the main objective of the genome scans is to detect true selective loci among the outliers. Our results suggest that a low critical probability ($\alpha =$

0.05%) is necessary (Figures 5–7) for this objective to be achieved successfully, but the method is not reliable for small values of \bar{s} or low proportions of selective loci in the genome. For example, for a proportion of selective loci of 1%, the percentage of selective loci present among the outliers is $\sim 10\text{--}20\%$ for $\bar{s} = 0.005$ and 40–60% for $\bar{s} = 0.5$ when the trimmed mean F_{ST} or the median F_{ST} are used (Figure 6). The use of the nontrimmed mean F_{ST} cannot be recommended in this context. For a proportion of selective loci between 1 and 5%, the percentage of selective loci present among the outliers is 100%, denoting a perfect success (Figure 6B). However, when the proportion of selective loci is 10%, no outliers are detected at all (Figure 6A) because the mean F_{ST} without any correction is too large.

BEAUMONT and BALDING (2004) simulations indicated an appreciable discrimination for selective loci when the selective effect is larger than about five times the migration rate. Thus, they found that $\sim 60\%$ of outliers were truly selective loci considering selection coefficients of $s = 0.02$, and this rate of success was improved up to 80–100% for selective effects between 0.05 and 0.2 (results deduced from their Table 3). This result is in agreement with our observations. For a neutral $F_{ST} = 0.025$ the migration rate was $m = 0.006$, so that selective effects larger than $5m = 0.03$ should be more likely detected. Figure 7H indicates that the percentage of selective loci in outliers is close to 100% for $\bar{s} = 0.05\text{--}0.5$, but $<30\%$ for $\bar{s} = 0.005$. Analogously, for a neutral $F_{ST} = 0.1$ the migration rate was $m = 0.00112$, so that selective effects >0.006 should be more likely detected. Figure 7K shows that the percentage of selective loci in outliers is $>70\%$ for $\bar{s} > 0.005$, but only $\sim 20\%$ for $\bar{s} = 0.001$. This agreement between both studies is very reassuring considering that the two models differ in a number of aspects: (i) BEAUMONT and BALDING (2004) analyzed codominant markers using the FDIST program whereas we studied dominant markers with the DFDIST software, (ii) the number of subpopulations considered was $n = 6$ whereas in our case $n = 2$, (iii) the selective coefficients were constant instead of variable, and (iv) the model of directional selection implied divergent selection whereas in our study only one of the two subpopulations was assumed to be under selection.

We also found that subpopulation sample size has some impact on the detection of selective loci, but this is limited beyond $M = 40$ individuals or so (Figure 7, M–O). This is in agreement with the conclusion of BEAUMONT and NICHOLS (1996) that there is little improvement in the performance of the FDIST method by having sample sizes larger than 25 diploids per subpopulation.

We can conclude that the detection of selective loci among outliers depends enormously on the conditions of the analysis and the proportion and effects of the selective loci. Using low critical probability threshold values (*e.g.*, $\alpha = 0.05\%$) and a trimmed or the median F_{ST}

seem the most appropriate procedures. We have not investigated the performance of an iterative process to obtain the mean F_{ST} of the null neutral model, suggested by BEAUMONT and NICHOLS (1996) (see BEAUMONT and BALDING 2004), whereby the raw mean F_{ST} is first calculated, then the outliers removed, and the process repeated until no outliers are detected (see, e.g., PAPA *et al.* 2007 for an empirical application). However, we have seen that a high degree of success is accomplished only when the proportion of selective loci in the genome is large and their average effect is also large. If these circumstances do not hold, the number of false positives can be very substantial and, therefore, the sequential removal of outliers is unlikely to work successfully. Another more reliable way to detect truly selective loci may be by combining the results of different tests (VASEMÄGI *et al.* 2005; BONIN *et al.* 2006), but this has yet to be checked by further analyses and simulations.

Impact of homoplasy on the detection of selective loci: The impact of homoplasy on the identification of selective loci was investigated assuming different numbers of loci and, consequently, of experimental fragments analyzed per primer combination (Figure 8). The main observation is that homoplasy causes a slight overestimation in the proportion of outliers and a more substantial underestimation in the proportion of selective loci detected. This latter can reach up to 15% when the number of experimental bands per primer combination is on the order of ≥ 75 –100 (Figure 8, C, F, and I). Thus, this aspect inherent to AFLP data should be taken into consideration in genomewide scans of selective loci.

Two additional important issues should be taken into account when using AFLP fragments in the detection of selective loci. First, in the cases of substantial homoplasy, some bands detected as outliers, even if they really correspond to true selective genes, do not necessarily carry a single selective fragment but they can also carry one or more neutral fragments. Therefore, a further experimental analysis of these bands should be carried out to disentangle selective and neutral comigrating fragments. The second issue is that the above threshold range refers to the number of experimental bands observed at the population level. If sample sizes are very small and the frequency of the allele producing the band is also low, the number of experimental bands observed can be much lower than the real one in the population, producing a false impression that the impact of homoplasy in the study will be low. For example, assuming genotypic frequencies at Hardy–Weinberg equilibrium and being \bar{p} the average frequency of the marker allele associated to the experimental bands, the ratio of the number of population bands to the number of observed bands is $1/[1 - (1 - \bar{p})^{2M}]$ if M individuals are sampled in the population. If this number is larger than 10 the ratio is close to one, but if few individuals are sampled the ratio could be very large. For example, if $M = 1$ (e.g., in phylogenetic studies), the observation of 60 bands per

primer combination corresponds, in fact, to 315.8 bands per primer combination in the population. Analogously, if $M = 5$, the observation of 60 bands in this sample corresponds to 92.1 bands per primer combination in the population. Therefore, the amount of homoplasy in these cases can be much larger than that inferred from the observed number of bands (60).

We thank Juan Galindo and two anonymous referees for useful comments on the manuscript. This work was funded by the Ministerio de Educación y Ciencia and Fondos Feder (CGL2006-13445-C02/BOS, CGL2004-03920/BOS), Plan Estratégico del Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (CPE03-004-C2), and Xunta de Galicia.

LITERATURE CITED

- ACHERÉ, V., J. M. FAVRE, G. NESNARD and S. JEANDROZ, 2005 Genomic organization of molecular differentiation in Norway spruce (*Picea abies*). *Mol. Ecol.* **14**: 3191–3201.
- AKASHI, H., R. M. KLIMAN and A. EYRE-WALKER, 1998 Mutation pressure, natural selection, and the evolution of base composition in *Drosophila*. *Genetica* **102/103**: 49–60.
- ALBACH, D. C., P. SCHÖNSWETTER and A. TRIBSCH, 2006 Comparative phylogeography of the *Veronica alpina* complex in Europe and North America. *Mol. Ecol.* **15**: 3269–3286.
- ALTHOFF, D. M., M. A. GITZENDANNER and K. A. SEGRAVES, 2007 The utility of amplified fragment length polymorphisms in phylogenetics: a comparison of homology within and between genomes. *Syst. Biol.* **56**: 477–484.
- BEAUMONT, M. A., 2005 Adaptation and speciation: What can F_{ST} tell us? *Trends Ecol. Evol.* **20**: 435–440.
- BEAUMONT, M. A., and D. J. BALDING, 2004 Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* **13**: 969–980.
- BEAUMONT, M. A., and R. A. NICHOLS, 1996 Evaluating loci for use in the genetic analysis of population structure. *Proc. R. Soc. Lond. B* **263**: 1619–1626.
- BENSCH, S., and M. AKESSON, 2005 Ten years of AFLP in ecology and evolution: Why so few animals? *Mol. Ecol.* **14**: 2899–2914.
- BONIN, A., P. TABERLET, C. MIAUD and F. POMPANON, 2006 Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Mol. Biol. Evol.* **23**: 773–783.
- BONIN, A., D. EHRLICH and S. MANEL, 2007 Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Mol. Ecol.* **16**: 3737–3758.
- BOWCOCK, A. M., J. R. KIDD, J. L. MOUNTAIN, J. M. HEBERT, L. CAROTENUTO *et al.*, 1991 Drift, admixture, and selection in human evolution. *Proc. Natl. Acad. Sci. USA* **88**: 839–843.
- BUSSEL, J. D., M. WAYCOTT and J. A. CHAPPILL, 2005 Arbitrarily amplified DNA markers for phylogenetic inference. *Perspect. Plant Ecol. Evol. Syst.* **7**: 3–26.
- CAMPBELL, D., and L. BERNATCHEZ, 2004 Generic scan using AFLP markers as a means to assess the role of directional selection in the divergence of sympatric whitefish ecotypes. *Mol. Biol. Evol.* **21**: 945–956.
- EL-RABEY, H. A., A. BADR, R. SCHAFFER-PREGL, W. MARTIN and F. SALAMINI, 2002 Speciation and species separation in *Hordeum L.* (*Poaceae*) resolved by discontinuous molecular markers. *Plant Biol.* **4**: 567–575.
- ENJALBERT, J., X. DUAN, M. LECONTE, M. S. HOVMØLLER and C. DE VALLAVIELLE-POPE, 2005 Genetic evidence of local adaptation of wheat yellow rust (*Puccinia striiformis* f. sp. *tritici*) within France. *Mol. Ecol.* **14**: 2065–2073.
- EXCOFFIER, L., J. NOVEMBRE and S. SCHNEIDER, 2000 SIMCOAL: a general coalescent program for the simulation of molecular data in interconnected populations with arbitrary demography. *J. Hered.* **91**: 506–509.

- FANG, J., J. TAO and C. T. CHAO, 2006 Genetic diversity in fruiting-mei, apricot, plum and peach revealed by AFLP analysis. *J. Hort. Sci. Biotechnol.* **81**: 898–902.
- FERNÁNDEZ, B., A. GARCÍA-DORADO and A. CABALLERO, 2005 The effect of antagonistic pleiotropy on the estimation of the average coefficient of dominance of deleterious mutations. *Genetics* **171**: 2097–2112.
- GORT, G., W. J. M. KOOPMAN and A. STEIN, 2006 Fragment length distributions and collision probabilities for AFLP markers. *Biometrics* **62**: 1107–1115.
- HANSEN, M., T. KRAFT, M. CHRISTIANSEN and N.-O. NILSSON, 1999 Evaluation of AFLP in *Beta*. *Theor. Appl. Genet.* **98**: 845–852.
- INNAN, H., R. TERAUCHI, G. KAHL and F. TAJIMA, 1999 A method for estimating nucleotide diversity from AFLP data. *Genetics* **151**: 1157–1164.
- JUMP, A. S., J. M. HUNT, J. A. MARTÍNEZ-IZQUIERDO and J. PEÑUELAS, 2006 Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. *Mol. Ecol.* **15**: 3469–3480.
- KALITA, M., and W. MALEK, 2006 Application of the AFLP method to differentiate *Genista tinctoria* microsymbionts. *J. Gen. Appl. Microbiol.* **52**: 321–328.
- KIDD, M. R., C. E. KIDD and T. D. KOCHER, 2006 Axes of differentiation in the bower-building cichlids of lake Malawi. *Mol. Ecol.* **15**: 459–478.
- KOOPMAN, W. J. M., and G. GORT, 2004 Significance tests and weighted values for AFLP similarities, based on *Arabidopsis in silico* AFLP fragment length distributions. *Genetics* **167**: 1915–1928.
- LAITUNG, B., E. CHAUVET, N. FEAU, K. FÈVE, L. CHIKHI *et al.*, 2004 Genetic diversity in *Trechochaetum elegans*, a mitosporic aquatic fungus. *Mol. Ecol.* **13**: 1679–1692.
- LEWONTIN, R. C., and J. KRAKAUER, 1973 Distribution of gene frequency as a test of the theory of the selected neutrality of polymorphism. *Genetics* **74**: 175–195.
- MEALOR, B. A., and A. L. HILD, 2006 Potential selection in native grass populations by exotic invasion. *Mol. Ecol.* **15**: 2291–2300.
- MEHANDA, S. M., B. R. BAUM, D. A. JOHNSON and J. T. ARNASON, 2004 Sequence assessment of comigrating AFLP™ bands in *Echinacea*—implications for comparative biological studies. *Genome* **47**: 15–25.
- MENDELSON, T. C., and K. L. SHAW, 2005 Use of AFLP markers in surveys of arthropod diversity. *Methods Enzymol.* **395**: 161–177.
- MEUDT, H. M., and A. C. CLARKE, 2007 Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends Plant Sci.* **12**: 106–117.
- MILLER, N. J., M. CIOSI, T. W. SAPPINGTON, S. T. RATCLIFFE, J. L. SPENCER *et al.*, 2007 Genome scan of *Diabrotica virgifera virgifera* for genetic variation associated with crop rotation tolerance. *J. Appl. Entomol.* **131**: 378–385.
- MUELLER, U. G., and L. L. WOLFENBARGER, 1999 AFLP genotyping and fingerprinting. *Trends Ecol. Evol.* **14**: 389–394.
- MURRAY, M. C., and M. P. HARE, 2006 A genome scan for divergent selection in a secondary contact zone between Atlantic and Gulf of Mexico oysters, *Crassostrea virginica*. *Mol. Ecol.* **15**: 4229–4242.
- NEI, M., 1973 Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* **70**: 3321–3323.
- NIELSEN, R., 2005 Molecular signatures of natural selection. *Annu. Rev. Genet.* **39**: 197–218.
- NOSIL, P., S. P. EGAN and D. J. FUNK, 2008 Heterogeneous genomic differentiation between walking-stick ecotypes: “isolation by adaptation” and multiple roles for divergent selection. *Evolution* **62**: 316–336.
- O’HANLON, P. C., and R. PEAKALI, 2000 A simple method for the detection of size homoplasy among amplified fragment length polymorphism fragments. *Mol. Ecol.* **9**: 815–816.
- PAPA, R., E. BELLUCCI, M. ROSSI, S. LEONARDI, D. RAU *et al.*, 2007 Tagging the signatures of domestication in common bean (*Phaseolus vulgaris*) by means of pooled DNA samples. *Ann. Botany* **100**: 11039–11051.
- PETERS, J. L., H. CONSTANDT, P. NEYT, G. CNOPS, J. ZETHOF *et al.*, 2001 A physical amplified fragment-length polymorphism map of *Arabidopsis*. *Plant Physiol.* **127**: 1575–1589.
- ROMBAUTS, S., Y. VAN DE PEER and P. ROUZÉ, 2003 AFLP in silico, simulating AFLP fingerprints. *Bioinformatics* **19**: 776–777.
- ROUPPE VAN DER VOORT, J. N., P. VAN ZANDVOORT, H. J. VAN ECK, R. T. FOLKERTSMA, R. C. HUTTEN *et al.*, 1997 Use of allele specificity of comigrating AFLP markers to align genetic maps from different potato genotypes. *Mol. Gen. Genet.* **255**: 438–447.
- SCHLÖTTERER, C., 2002 A microsatellite-based multilocus screen for the identification of local selective sweeps. *Genetics* **160**: 753–763.
- SCOTTI-SAINTAGNE, C., S. MARIETTE, I. PORTH, P. G. GOICOECHEA, T. BARRENECHE *et al.*, 2004 Genome scanning for interspecific differentiation between two closely related oak species [*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.]. *Genetics* **168**: 1615–1626.
- SMITH, T. B., B. MILÁ, G. F. GREYER, H. SLABBEOORN, I. SEPIL *et al.*, 2008 Evolutionary consequences of human disturbance in a rainforest bird species from Central Africa. *Mol. Ecol.* **17**: 58–71.
- SNÄLL, T., J. FOGELQVIST, P. J. RIBEIRO and M. LASCoux, 2004 Spatial genetic structure in two congeneric epiphytes with different dispersal strategies analysed by three different methods. *Mol. Ecol.* **13**: 2109–2119.
- STEIN, L. D., Z. BAO, D. BLASIAK, T. BLUMENTHAL, M. R. BRENT *et al.*, 2003 The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol.* **1**: 166–192.
- STORZ, J. F., 2005 Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol. Ecol.* **14**: 671–688.
- STORZ, J. F., and J. M. DUBACH, 2004 Natural selection drives altitudinal divergence at the albumin locus in deer mice, *Peromyscus maniculatus*. *Evolution* **58**: 1342–1352.
- STORZ, J. F., and M. W. NACHMAN, 2003 Natural selection on protein polymorphism in the rodent genus *peromyscus*: evidence from interlocus contrasts. *Evolution* **57**: 2628–2635.
- STORZ, J. F., B. A. PAYSEUR and M. W. NACHMAN, 2004 Genome scans of DNA variability in human reveal evidence for selective sweeps outside of Africa. *Mol. Biol. Evol.* **21**: 1800–1811.
- TROELL, K., A. ENGSTROM, D. A. MORRISON, J. G. MATSSON and J. HOGLUND, 2006 Global patterns reveal strong population structure in *Haemonchus contortus*, a nematode parasite of domesticated ruminants. *Int. J. Parasitol.* **36**: 1305–1316.
- VASEMÄGI, A., J. NILSSON and C. R. PRIMMER, 2005 Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). *Mol. Biol. Evol.* **22**: 1067–1076.
- VEKEMANS, X., T. BEAUWENS, M. LEMAIRE and I. ROLDÁN-RUIZ, 2002 Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Mol. Ecol.* **11**: 139–151.
- VITALIS, R., K. DAWSON and P. BOURSOT, 2001 Interpretation of variation across loci as evidence of selection. *Genetics* **158**: 1811–1823.
- VOS, P., R. HOGERS, M. BLEEKER, T. VAN DE LEE, M. HORNES *et al.*, 1995 AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **21**: 4407–4414.
- WEIR, B., and C. COCKERHAM, 1984 Estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.
- WILDING, C. S., R. K. BUTLIN and J. GRAHAME, 2001 Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *J. Evol. Biol.* **14**: 611–619.
- WRIGHT, S., 1937 The distribution of gene frequencies in populations. *Proc. Natl. Acad. Sci. USA* **23**: 307–320.
- WRIGHT, S., 1951 The genetic structure of populations. *Ann. Eugen.* **15**: 323–354.
- YU, J., S. HU, J. WANG, G. K.-S. WONG, S. LI *et al.*, 2002 A draft sequence of the rice genome (*Oryza sativa* L. ssp. Indica). *Science* **296**: 79–92.
- ZHIVOTOVSKY, L. A., 1999 Estimating population structure in diploids with multilocus dominant DNA markers. *Mol. Ecol.* **8**: 907–913.