Microsatellite Allele Sizes: A Simple Test to Assess Their Significance on Genetic Differentiation

Olivier J. Hardy,*,[1] Nathalie Charbonnel,[†] Hélène Fréville[‡] and Myriam Heuertz*,[§]

**Laboratoire de Génétique et Ecologie Végétales, Université Libre de Bruxelles, 1160 Brussels, Belgium, [†]CEFE-CNRS, 34293 Montpellier Cedex 5, France, [‡]Department of Biology, Open University, Milton Keynes MK7 6AA, United Kingdom and [§]CRP-Gabriel Lippmann, CREBS Research Unit, 1511 Luxembourg, Luxembourg*

Manuscript received May 10, 2002
Accepted for publication December 10, 2002

ABSTRACT

The mutation process at microsatellite loci typically occurs at high rates and with stepwise changes in allele sizes, features that may introduce bias when using classical measures of population differentiation based on allele identity (*e.g.*, F_{ST}, Nei's *Ds* genetic distance). Allele size-based measures of differentiation, assuming a stepwise mutation process [*e.g.*, Slatkin's R_{ST}, Goldstein *et al.*'s $(\delta\mu)^2$], may better reflect differentiation at microsatellite loci, but they suffer high sampling variance. The relative efficiency of allele size- *vs.* allele identity-based statistics depends on the relative contributions of mutations *vs.* drift to population differentiation. We present a simple test based on a randomization procedure of allele sizes to determine whether stepwise-like mutations contributed to genetic differentiation. This test can be applied to any microsatellite data set designed to assess population differentiation and can be interpreted as testing whether $F_{ST} = R_{ST}$. Computer simulations show that the test efficiently identifies which of F_{ST} or R_{ST} estimates has the lowest mean square error. A significant test, implying that R_{ST} performs better than F_{ST}, is obtained when the mutation rate, μ, for a stepwise mutation process is (a) $\geq m$ in an island model (m being the migration rate among populations) or (b) $\geq 1/t$ in the case of isolated populations (t being the number of generations since population divergence). The test also informs on the efficiency of other statistics used in phylogenetical reconstruction [*e.g.*, *Ds* and $(\delta\mu)^2$], a nonsignificant test meaning that allele identity-based statistics perform better than allele size-based ones. This test can also provide insights into the evolutionary history of populations, revealing, for example, phylogeographic patterns, as illustrated by applying it on three published data sets.

MICROSATELLITE genetic markers—also called short tandem repeats (STRs) or simple sequence repeats (SSRs) because their polymorphism is based on the variation in the number of repeats of a simple DNA sequence (2–6 bases long)—are nowadays a tool of choice to address population genetics and demographic questions (*e.g.*, Estoup and Angers 1998).

Microsatellite loci are typically characterized by high mutation rates and hence a high level of polymorphism as well as by a mutation process that causes preferentially stepwise changes of the number of repeats [stepwise mutation model (SMM), Table 1] and thus allele size (*e.g.*, Zhu *et al.* 2000). Hence, the difference in size between two different alleles might be informative: The larger the difference, the higher the number of mutation events (thus time lapse) is expected to have occurred since common ancestry. There is thus a "memory" of past mutation events. Slatkin (1995) showed that if the mutational process follows a SMM, the expected squared difference between allele sizes is a linear function of the expected coalescence time of the alleles compared. On the contrary, if mutations result in one of K possible alleles at random [K-allele model (KAM), infinite-allele model (IAM); Table 1], comparison between any two different alleles (alleles not identical in state) bears the same information: At least one mutation has occurred since common ancestry; the mutation process is memoryless. Comparison of microsatellite alleles can thus provide two kinds of information: allele identity/nonidentity and allele size differences (throughout this article, allele identity refers to identity in state and not identity by descent).

Most statistics that describe genetic differentiation from genetic markers (*e.g.*, *F*-statistics) rely solely on allele identity information. This information is often used to infer phylogenetic relationships or to obtain indirect estimates of gene flow. In the first case, studied populations are assumed to have diverged by drift and mutation without gene flow, so that genetic differentiation informs on the time since the beginning of divergence (*e.g.*, Nei 1972). In the second case, studied populations are assumed to have diverged by drift up to a migration-drift equilibrium, so that genetic differentiation informs on the balance between drift and gene flow (*e.g.*, Slatkin 1985). For example, considering an island model of diploid populations (*i.e.*, a large number

[1]*Corresponding author:* Laboratoire de Génétique et Ecologie Végétales, Chaussée de Wavre 1850, B-1160 Brussels, Belgium.
E-mail: ohardy@ulb.ac.be

TABLE 1

Mutation models

Models	Effect of a mutation event
IAM: infinite-allele model	New allele (never observed previously) created
KAM: *K*-allele model	Mutation toward one of *K* possible allelic states (excluding the original state)
SMM: stepwise mutation model	Allele size increased or decreased by just 1 unit
GSM: generalized stepwise model	Allele size modified by *x* units, *x* being a random variable following any distribution of finite variance

IAM is a particular case of KAM for K = infinity. SMM is a particular case of GSM with $x = -1$ or 1. For SMM and GSM, the range of potential allele sizes is unbounded and the mutation rate is independent of prior allele size (Kimmel and Chakraborty 1996; Feldman *et al.* 1999).

of populations of effective size N receiving each generation a proportion m of genes taken randomly from the other populations) at migration-drift equilibrium, a commonly used relationship is $F_{ST} \approx 1/(1 + 4Nm)$ (Wright 1965). F_{ST} is a parameter describing the degree of genetic differentiation among populations and is defined as the correlation of allelic states between genes sampled within populations or, equivalently, $F_{ST} \equiv (Q_w - Q_b)/(1 - Q_b)$, where Q_w (Q_b) is the probability that two genes from the same population (different populations) are identical in state (Excoffier 2001). The product Nm, a demographic parameter describing the effective number of migrants per population and generation (gene flow), can thus be inferred from F_{ST}. Among other assumptions (*e.g.*, Whitlock and McCauley 1999), this relationship assumes a low mutation rate μ (notably $\mu \ll m$); otherwise $F_{ST} \approx 1/(1 + 4N(m + \mu))$ (Crow and Aoki 1984), and gene flow cannot be inferred from an estimate of F_{ST} unless μ is accurately known. As microsatellites typically have high μ (of the order of 10^{-5} to 10^{-2}; Jarne and Lagoda 1996), their use might lead to significantly biased gene dispersal estimates. Therefore, it has been argued that microsatellites are not adequate for large-scale studies of population genetic structure (*i.e.*, when m is likely to be very low and divergence time long) or that only loci with an intermediate level of polymorphism (suggesting moderate mutation rates) should be considered (Jarne and Lagoda 1996; Estoup and Angers 1998).

Alternative solutions to this problem have been proposed using statistics accounting for allele size information, such as *R*-statistics (Slatkin 1995; Rousset 1996; see also Balloux and Lugon-Moulin 2002 for a general discussion on *F*- and *R*-statistics when assessing population differentiation with microsatellites). Indeed, R_{ST} is an analog of F_{ST} based on allele size differences: It is a parameter defined as the correlation of allele sizes (rather than allelic states) between genes sampled within populations or, equivalently, $R_{ST} \equiv (S_b - S_w)/S_b$, where S_w (S_b) is the mean square difference in allele size for two genes from the same population (different populations; Excoffier 2001, a definition slightly different from Slatkin 1995). The analogy between the mathematical definitions of F_{ST} and R_{ST} is more obvious when noting that $(1 - Q)$ and S both express a degree of genetic variability, F_{ST} and R_{ST} expressing the proportion of variability that can be attributed to differentiation among populations. R_{ST} is related to gene flow in a way equivalent to F_{ST} [*e.g.*, $R_{ST} \approx 1/(1 + 4Nm)$ in an island model] but without assumption on the mutation rate so that, contrary to F_{ST}, the relationship remains valid for $\mu \geq m$ in an island model (Rousset 1996). Here, however, the mutation process is assumed to follow a pure SMM or a generalized stepwise model (GSM; Table 1). Allele size information is also exploited by several measures of genetic distances developed for phylogenetic reconstruction (*e.g.*, Goldstein *et al.* 1995b; Shriver *et al.* 1995; Kimmel *et al.* 1996), assuming also a SMM or a GSM. There are, however, two important drawbacks when using allele size-based statistics. First, microsatellite mutations are known to deviate more or less strongly from an ideal SMM or GSM (reviewed in Estoup and Angers 1998; Ellegren 2000; Xu *et al.* 2000). These deviations can result in strongly biased estimates of divergence time or R_{ST}-based estimates of gene flow. Second, statistics based on allele size typically suffer high sampling variances when compared to their counterparts based on allele identity information (Goldstein *et al.* 1995b; Takezaki and Nei 1996), as was shown for R_{ST} and F_{ST} estimators (Slatkin 1995; Gaggiotti *et al.* 1999; Balloux and Goudet 2002). (As we are not dealing with the problematics of parameter estimation, we do not use different notations to distinguish F_{ST} and R_{ST} parameters from their respective estimators. In the following, F_{ST} and R_{ST} refer to estimators that are specified more accurately later on.)

On the basis of simulation results, Gaggiotti *et al.* (1999) suggested that for most typical sample sizes and genetic parameters encountered in experimental studies, F_{ST} should be preferred over R_{ST} to estimate gene flow parameters with microsatellites because it generally gave a lower mean square error (a measure of error accounting for both the bias and the standard error of the estimates) of Nm estimates. A similar study by

Balloux and Goudet (2002) showed that F_{ST} is more efficient in the case of high levels of gene flow whereas R_{ST} better reflects population differentiation under low gene flow. From simple theoretical considerations, one can predict that there is no gain in using R_{ST} over F_{ST} when $\mu \ll m$, as both would share identical expectations (Slatkin 1995; Rousset 1996), but F_{ST} should be preferred because of its lower standard error. However, it is difficult to know *a priori* which conditions apply for a given data set and thus to determine which statistic is the most appropriate.

Comparing F_{ST} and R_{ST} values computed on the same data can provide valuable insights into the main causes of population differentiation, *i.e.*, drift *vs.* mutation, because these statistics share equal expectations when differentiation is caused solely by drift, whereas R_{ST} is expected to be larger than F_{ST} under a contribution of stepwise-like mutations (*e.g.*, Michalakis and Veuille 1996; Ross *et al.* 1997; Estoup *et al.* 1998; Lugon-Moulin *et al.* 1999). Their comparison can reveal phylogeographic patterns, that is, when genetic divergence between distinct alleles is related to geographical separation. However, no procedure has been developed to date for testing whether single-locus R_{ST} and F_{ST} estimates are significantly different.

This article proposes a simple testing procedure based on allele size randomizations to determine if mutations following a SMM-like process contribute to genetic differentiation. The test can reveal whether allele identity-based or allele size-based statistics should be most adequate to analyze microsatellite data sets. A nonsignificant test suggests then that F_{ST} should be preferred over R_{ST} or, more generally, that statistics based on allele identity are likely to perform better than counterparts based on allele size information. When mutations are known to follow a SMM-like process, the test can also assess the relative importance of the mutation rates *vs.* the migration rate or *vs.* the reciprocal of the divergence time in the case of isolated populations. This procedure can be interpreted as testing whether $R_{ST} = F_{ST}$ and could therefore be used to reveal phylogeographic patterns.

In the following, we present the test, validate it by simulations, explore its power in different contexts by simulations again, and apply it on three data sets from published experimental studies. Emphasis is given to the usefulness of the test to determine the efficiency of F_{ST} *vs.* R_{ST} for inferential purposes. Its usefulness to assess the efficiency of other statistics based on allele identity *vs.* allele size is addressed in the discussion, together with other potential applications.

A SIMPLE TEST ON ALLELE SIZE INFORMATION CONTENT

The test indicates whether allele sizes provide information on population differentiation given a data set, that is, whether shifts in allele sizes resulting from stepwise-like mutations contribute to population differentiation. Contribution of stepwise-like mutations to genetic differentiation requires (1) that the mutation process is at least partially SMM-like and (2) that the mutation rate, μ, is large enough relative to the effect of drift and migration (*e.g.*, $\mu \geq m$; otherwise new mutations are quickly spread beyond their native population by migration). Table 2 outlines the null hypotheses that can be tested, presenting a general null hypothesis as well as specific null hypotheses holding under particular prior assumptions.

The principle of the test is based on obtaining a distribution of a statistic under the null hypothesis (H_0) that differences in allele sizes do not contribute to population differentiation. Therefore, we use a randomization procedure whereby the different allele sizes observed at a locus for a given data set are randomly permuted among allelic states. To better figure out the procedure, one may dissociate allelic state, identified, for example, by a letter (*e.g.*, *a*, *b*, *c*, *d*, and *e* if there are five different alleles), and allele size, identified by a number (*e.g.*, 4, 5, 7, 8, and 11, each representing the number of sequence repeats), given that there is a one-to-one correspondence between allelic state and allele size. Before randomization, the allele size attributed to each allelic state is the actual allele size (*e.g.*, *a*, 4; *b*, 5; *c*, 7; *d*, 8; and *e*, 11). Throughout the randomization procedure, genotypes are defined in terms of allelic states and are not modified, but allele sizes are randomly reassigned among allelic states (*e.g.*, *a*, 7; *b*, 4; *c*, 11; *d*, 5; and *e*, 8). After such a randomization, any two genes originally having the same allele size remain identical, although it can be for another allele size, whereas any two genes originally bearing different alleles of small size difference may bear alleles of large size difference, or reciprocally. Hence, the allele identity information is kept intact but not the allele size information. Under the null hypothesis (Table 2, case 1), the randomization procedure should not affect the expectation of a measure of differentiation such as R_{ST}. On the contrary, if allele sizes contribute to genetic differentiation, the R_{ST} computed after allele size permutation (hereafter called pR_{ST}) would depend solely on allele identity/nonidentity and hence have a smaller expectation than the value computed before randomization. The test can thus be designed by comparing the observed R_{ST} value (before randomization) to the distribution of pR_{ST} values obtained for all possible configurations of allele size permutations (or a representative subset of them, as the total number of different configurations quickly becomes enormous when the number of alleles exceeds 7 or 8). From this comparison, a probability that the null hypothesis holds can be estimated as the proportion of pR_{ST} values larger than the observed R_{ST} (one-tailed test). Note that the mean pR_{ST} should equal in expectation the F_{ST} computed on the same data (not accounting for potential statistical bias), as is confirmed later.

On a single locus, such a test can be applied only if a sufficient number of different alleles (n) are in the

TABLE 2

Hypotheses tested by allele size permutations applied on R_{ST}

Null hypothesis H_0	Alternative hypothesis H_a
1. General hypotheses	
No contribution of stepwise mutations to genetic differentiation	Contribution of stepwise mutations to genetic differentiation
$R_{ST} = F_{ST}$	$R_{ST} > F_{ST}$
2. Specific hypotheses when stepwise mutations occurred	
Mutations negligible relative to drift	Mutations not negligible relative to drift
a. Island model:	
$\mu \ll m$	$\mu \geq m$
b. Isolated population model:	
$\mu \ll 1/t$	$\mu \geq 1/t$
3. Specific hypotheses when mutations contributed to genetic differentiation (*e.g.*, $\mu \geq m$, $\mu \geq 1/t$)	
No stepwise-like mutations	Stepwise-like mutations
KAM, IAM	*e.g.*, SMM, GSM

μ is the mutation rate, m is the migration rate per generation, t is the divergence time in number of generations since populations' isolation. The sign "≥" should be understood as "larger than or in the same order of magnitude as." The neutrality of the markers is assumed throughout. The general hypotheses always hold whereas the specific hypotheses are context dependent.

data set, as the number of different permutation configurations is equal to $n!$. Hence, five alleles (120 different configurations) appear to be a minimum to carry out such test at a type I error rate criterion of 5 or 1%. On a multilocus R_{ST} estimate, the test can be carried out by permuting allele sizes within each locus. It is noteworthy that the test makes no assumptions on the mutation model: A significant result (R_{ST} significantly $> pR_{ST}$) suggests that mutations contributed to genetic differentiation (*e.g.*, because $\mu \geq m$ in an island model) *and* that the mutation process follows at least partially a SMM (the test remains valid under deviations from the SMM). Neutrality with respect to natural selection is, however, assumed. When the test is significant, F_{ST} is likely to provide a biased estimate of gene flow parameters, but it cannot be concluded *a priori* that R_{ST} would necessarily perform better given its larger variance (which is even more pronounced when mutations of more than one step can occur; Zhivotovsky and Feldman 1995) and given the bias it may suffer when the mutation process deviates from the assumptions of the GSM (Estoup and Angers 1998). A nonsignificant result (R_{ST} not significantly different from pR_{ST}) would suggest that allele size is not informative for population differentiation, because the mutation process is not stepwise-like and/or because mutations had not contributed to differentiation (*e.g.*, because $\mu \ll m$ in an island model). In this case, F_{ST} should surely be preferred over R_{ST} (although it would not ensure that F_{ST} provides a correct estimate of gene flow given the many other sources of bias related to population models; Whitlock and McCauley 1999).

Which hypotheses can be tested and with which statistics? Simulations permit validation of the allele size permutation test and assess its power. But it is first necessary to insist on what can be tested (Table 2).

Randomizing allele sizes creates replicates of a data set for a mutation process following a KAM (or IAM) because, under this model, allele size is irrelevant and interchanging them is like replicating the past mutation processes leading to the present data set but with other randomly chosen alleles after each mutational event. Hence, one possible application of the allele size randomization procedure is to test whether the mutation process follows a KAM (Table 2, case 3). For this purpose, randomizing allele sizes can be applied on any statistic based on allele size, not only *R*-statistics but also various genetic distances for stepwise mutation models such as $(\delta\mu)^2$ (*e.g.*, Goldstein *et al.* 1995b; Shriver *et al.* 1995), or simply on the total variance in allele size. It is, however, already well established that the large majority of microsatellite loci do not conform to a KAM, and the interesting question about the mutation process of microsatellites is rather how it deviates from an ideal SMM (Estoup and Angers 1998). Therefore, using the allele size permutation procedure to test for the KAM is not discussed further.

A second application of the allele size permutation procedure, here assuming *a priori* that mutations follow at least partially a SMM-like process, is to test whether mutation has contributed to population divergence (Table 2, case 2). In other words, we can test whether the migration rate (m) among populations, or the reciprocal of the number of generations (t) since population divergence, is large compared to the mutation rates ($\mu \ll m$ or $\mu \ll 1/t$, respectively; Table 2, cases 2a and 2b). The allele size permutation test is the most interesting to address this question, because there is enough

evidence that most microsatellites follow a SMM-like process (*e.g.*, Ellegren 2000; Xu *et al.* 2000; Zhu *et al.* 2000; Renwick *et al.* 2001). However, for this purpose, allele size permutation cannot be applied to any statistic based on allele size: It performs well on *R*-statistics, which are ratios of allele size variance components, but not on genetic distances such as the Goldstein *et al.* (1995a) $(\delta\mu)^2$ statistic, which is a between-populations component of allele size variance. The reason is that random permutations of allele sizes not only remove the within-population covariance between allele sizes for different alleles, but also modify the allele size variance under SMM or GSM, because the expected frequency distribution of allele sizes is not uniform (Donnelly 1999). Statistics expressing a component of allele size variance, such as the $(\delta\mu)^2$ statistic, will always be affected by a change of the allele size variance, no matter whether or not mutations contributed to differentiation. On the contrary, statistics based on a ratio of variance components, such as R_{ST}, will not be affected if the within- and among-populations components of variance are multiplied by factors having the same expectations. The simulations presented hereafter show that this is what occurs when there is no within-population covariance between allele sizes for different alleles (*i.e.*, differentiation due to drift and not stepwise mutations).

To show that the allele size permutation test is adequate for the R_{ST} statistic but not the $(\delta\mu)^2$ statistic when testing $m \gg \mu$ or $1/t \gg \mu$ (under the *a priori* assumption that the mutation process is stepwise-like; Table 2, cases 2), we simulated a random-mating population of diploid individuals (population size $N = 1000$ individuals) at mutation-drift equilibrium ($\mu = 0.001$) under the SMM. The allele size permutation test (1000 randomizations) was then applied on R_{ST} and $(\delta\mu)^2$ computed between two independent samples (sample size $n = 100$ individuals) from that population for each of 200 simulated loci (the two samples thus represent undifferentiated subpopulations). The computer programs used for simulations and computations are described below. We report the percentage of loci for which the tests were significant (%RHo) according to the type I error rate criterion (α, the probability of rejecting the null hypothesis when it is true). Because the null hypothesis to be tested ($1/t \gg \mu$) is met by simulations, a valid testing procedure must ensure that %RHo = α; otherwise it means that the procedure is not adequate to test this null hypothesis. Figure 1 shows that the allele size randomization testing procedure is indeed valid when applied on R_{ST} but not on $(\delta\mu)^2$.

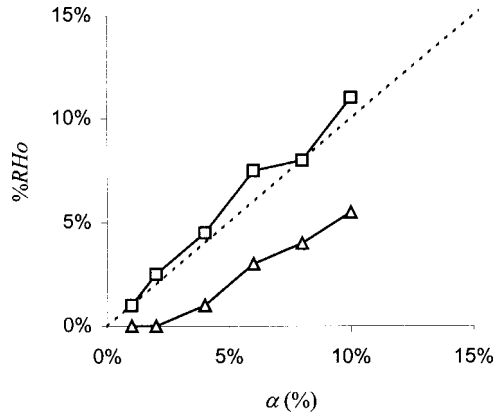

Figure 1.—Control of the validity of the allele size permutation test when applied on R_{ST} (□) or $(\delta\mu)^2$ (△) statistics computed between two samples from a population at mutation-drift equilibrium under the SMM. The percentage of loci with the null hypothesis rejected (%RHo) is shown as a function of the type I error rate criterion (α), and the dashed line shows the %RHo = α relationship expected under the null hypothesis for a valid testing procedure. The null hypothesis of interest is whether the mutation rate is negligible, given that the mutation process is stepwise-like (Table 2, case 2). Results show that the allele size permutation procedure applied on $(\delta\mu)^2$ is not suited to test this hypothesis.

Power of the test under SMM: To investigate the power of the test when testing if mutations contributed to population differentiation under the SMM (Table 2, cases 2), we checked the procedure on artificial data sets with realistic sample sizes derived from Monte Carlo simulations of populations made of diploid hermaphrodites. Three sets of demographic situations were simulated: (1) an island model at drift-migration-mutation equilibrium, (2) a model of two isolated populations having diverged from a common ancestral population at mutation-drift equilibrium, and (3) a linear stepping-stone model (gene flow restricted to adjacent populations) at drift-migration-mutation equilibrium. The island model was composed of 10 populations, consisting of 100 individuals each, and new generations were obtained by drawing genes at random from the population with probability $1 - m$ or from the other populations with probability m. The isolated population model was composed of two random-mating populations, consisting of 500 individuals each, and having diverged for t generations. The stepping-stone model was composed of 30 aligned populations, consisting of 50 individuals each, and new generations were obtained by drawing genes at random from the population with probability $1 - m$ or from the two adjacent populations with probability m.

The genetic parameters simulated were the following: At the initial stage all populations were fixed for one allele; 10 loci were simulated with mutations following a SMM and $\mu = 10^{-3}$ at all loci without size constraints. Simulations were run for a sufficient time to reach a steady state for total- and within-population gene diversity parameters, and then a sample of individuals representative of common experimental studies was extracted and analyzed. To obtain accurate estimates, 200 replicates were run for each set of conditions. Simulations were carried out using the software EASYPOP ver. 1.7.4 (Balloux 2001). Allele size permutation tests (with

1000 randomizations) and computations of F_{ST} and R_{ST} on the samples extracted were done with the program SPAGeDi (Hardy and Vekemans 2002). Single-locus and multilocus F_{ST} and R_{ST} were estimated following Weir and Cockerham (1984) and Michalakis and Excoffier (1996), respectively. It should be noted that this R_{ST} (an estimator of the parameter called ρ_{ST} by Rousset 1996) differs somewhat from Slatkin's (1995) original definition (Michalakis and Excoffier 1996) but is better suited for comparison with the F_{ST} estimator of Weir and Cockerham (1984) (called θ by these authors) and for demographic parameter estimations (Rousset 1996). Both these F_{ST} and R_{ST} estimators proceed by a standard hierarchical ANOVA where the observed variance (σ^2) of allele identity per locus and per allele (F_{ST}), or the variance of allele size per locus (R_{ST}), is partitioned into three components (random effects): among populations (σ_a^2), among individuals within population (σ_b^2), and between genes within individual within population (σ_c^2). F_{ST} and R_{ST} are then estimated as $\sigma_a^2/(\sigma_a^2 + \sigma_b^2 + \sigma_c^2)$ (single-locus R_{ST}) or $\Sigma\sigma_a^2/\Sigma(\sigma_a^2 + \sigma_b^2 + \sigma_c^2)$, where the summations apply over all loci (multilocus R_{ST}), all alleles of a locus (single-locus F_{ST}), or all alleles and loci (multilocus F_{ST}; Excoffier 2001).

For the island model, simulations were run for 5000 generations with migration rates among populations varying from 10^{-4} to 10^{-1} (*i.e.*, $m = 0.1$–100μ) according to the runs. Global R_{ST}, F_{ST}, and pR_{ST} (for 1000 randomizations) were computed on a total sample of 300 individuals (30 individuals from each population). For the isolated populations model, a single population of 1000 individuals was simulated for 5000 generations, and then it was divided into two isolated subpopulations of 500 individuals that were run for 30–10,000 additional generations (*i.e.*, $1/t = 0.1$–33μ). R_{ST}, F_{ST}, and pR_{ST} (for 1000 randomizations) were computed on a total sample of 100 individuals (50 individuals from each subpopulation). For the stepping-stone model, 10,000 generations were simulated with a migration rate of 0.1 (0.05 between any two adjacent populations). Analyses were carried out on a sample of 20 individuals from each of the 30 populations (total sample size of 600 individuals). Pairwise $F_{ST}/(1 - F_{ST})$ and $R_{ST}/(1 - R_{ST})$ ratios were computed for each pair of populations, and these values were averaged over all pairs separated by 1, 2, 3, . . . , 20 steps (20 distance classes). Allele size permutation tests were applied on averaged pairwise $R_{ST}/(1 - R_{ST})$ ratios per distance class to provide $pR_{ST}/(1 - pR_{ST})$ values per distance class (1000 permutations). Here, pairwise $F_{ST}/(1 - F_{ST})$ and $R_{ST}/(1 - R_{ST})$ ratios were computed because theory predicts an approximate linear relationship with the linear distance between populations in one-dimensional isolation-by-distance models (Rousset 1997).

The validity of some of the simulation results could be verified by comparing them to theoretical expectations. For example, after 5000 generations of simulation of a single population of $N = 1000$ individuals (for the isolated population model), the average heterozygosity and average variance of allele size were equal to He = 0.68 and $V = 1.96$, respectively, with a mean number of alleles per locus of 5.8 (range, 3–11 alleles). These values are close to their expectations at mutation-drift equilibrium (Estoup and Cornuet 1999): Under strict SMM, $He = 1 - (1 + 8N\mu)^{-0.5} = 0.67$ and $V = 2N\mu = 2$. In the island model with 10 populations of 100 individuals each ($d = 10$, $N = 100$), average R_{ST} values were equal to 0.019, 0.197, 0.677, and 0.924 for $m = 10^{-1}$, 10^{-2}, 10^{-3}, and 10^{-4}, respectively (Figure 2A), in agreement with the expected values approximately equal to $1/(1 + 4Nm\ d/(d - 1)) = 0.022$, 0.184, 0.692, and 0.957, respectively (Rousset 1996). In the isolated populations model ($N = 500$), divergence time t can be estimated from the relationship $R_{ST}/(1 - R_{ST}) = t/2N$ (Slatkin 1995; Rousset 1996), giving estimates of $t =$ 97, 1132, and 11,301 for actual values of 100, 1000, and 10,000 generations, respectively. Finally, in the linear stepping-stone model ($N = 50$, $m = 0.1$), pairwise $R_{ST}/(1 - R_{ST})$ values increased linearly with the distance between populations (Figure 2C), giving a regression slope equal to 0.054, in agreement with the approximate expected value $1/(4Nm) = 0.050$ for the linear stepping-stone model (Rousset 1997).

Results from all simulations confirm that mean pR_{ST} values (*i.e.*, mean value computed after random permutations of allele size) are very close, though not exactly equal, to the F_{ST} values (Figure 2). For example, in the island model, the mean and standard deviation of the difference between F_{ST} and mean pR_{ST} values per locus were equal to 0.003 ± 0.007, 0.008 ± 0.012, and 0.010 ± 0.110 for $m = 10^{-2}$, 10^{-3}, and 10^{-4}, respectively. Hence, mean pR_{ST} values were on average slightly lower than F_{ST} values although, for a given locus, the difference between the two could be quite substantial, especially under very low migration rates. For the other simulations, mean pR_{ST} values were generally slightly higher than F_{ST} (Figure 2, B and C). We also observed that the discrepancy between F_{ST} and mean pR_{ST} was much lower for multilocus than for single-locus estimates.

As expected, R_{ST} values are similar to F_{ST} values whenever $m \gg \mu = 0.001$ (island model), $1/t \gg \mu$ (diverging populations model), or populations are close (stepping-stone model with $m \gg \mu$). On the contrary, R_{ST} becomes considerably larger than F_{ST} when $m \leq \mu$ (island model), $1/t \leq \mu$ (diverging populations model), or when populations are separated by more than five steps (stepping-stone model; Figure 2).

To assess the power of the allele size permutation test, we present in Figure 2 (graphs on the right) the percentage of statistically significant tests (%RHo) among 200 simulation replicates (using $\alpha = 5\%$) according to (1) the migration rate m (island model), (2) the divergence time t in number of generations since isolation (isolated two-population model), and (3) the

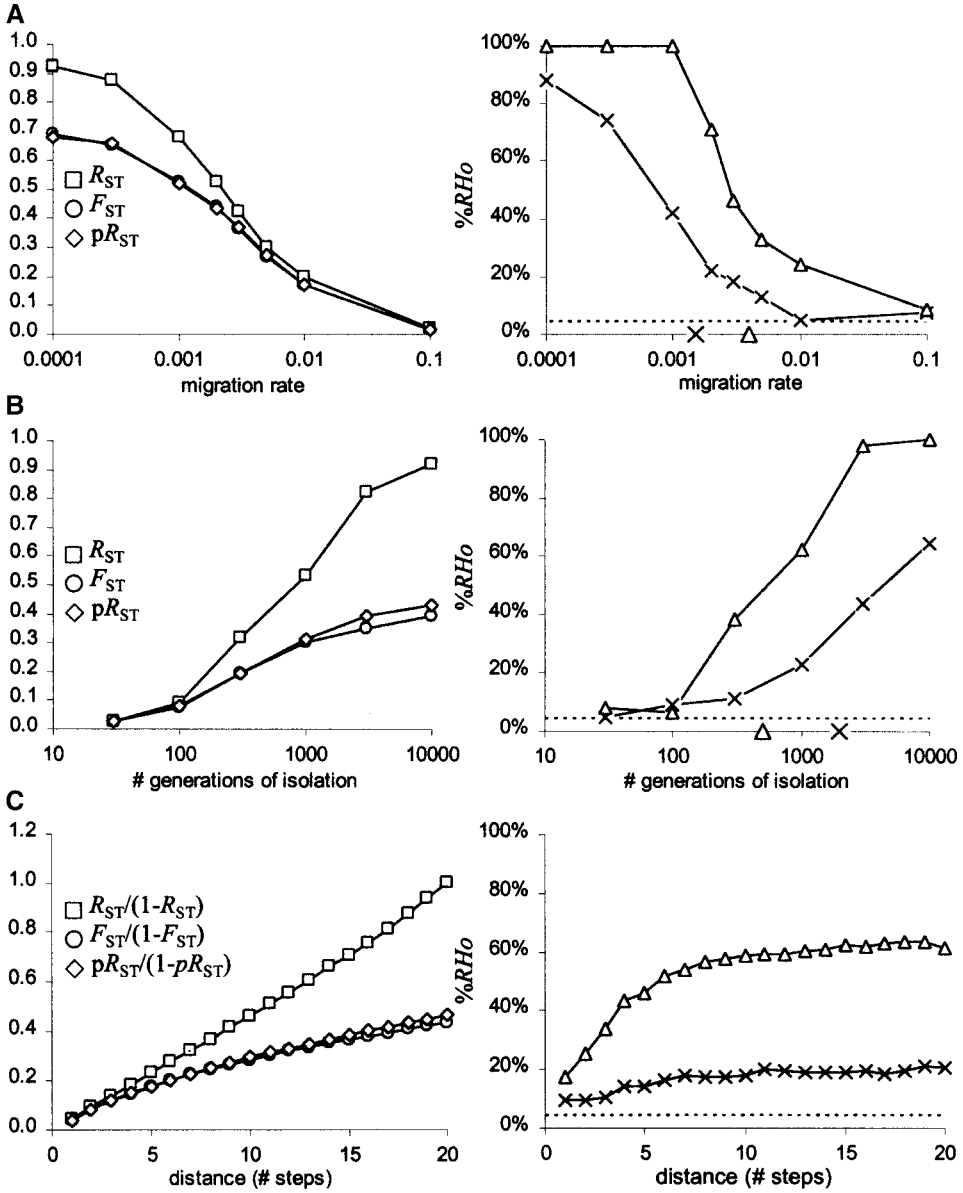

FIGURE 2.—Simulation results for (1) an island model with migration rate m (A), (2) a two-population model isolated for t generations (B), and (3) a linear stepping-stone model of 30 populations (C). Graphs on the left show R_{ST} (□), F_{ST} (○), and mean pR_{ST} (◇) values (mean multilocus estimates based on 10 loci and 200 replicates) according to m (A), t (B), or the number of steps separating populations (C). In C, averaged pairwise $R_{ST}/(1 - R_{ST})$, $F_{ST}/(1 - F_{ST})$, and mean $pR_{ST}/(1 - pR_{ST})$ ratios over all pairs separated by given numbers of steps are represented. Graphs on the right illustrate the power of the allele size permutation tests by giving the percentages of significant tests (%RHo) on R_{ST} estimates [or average pairwise $R_{ST}/(1 - R_{ST})$ ratios] based on a single locus (×) or 10 loci (△) (*i.e.*, multilocus estimate) and considering a type I error rate criterion α of 5% (dotted line). The symbols (× and △) on the horizontal axes of graphs A and B show the values at which the mean square errors of F_{ST} and R_{ST} are approximately equal.

distance d in number of steps between populations (stepping-stone model). This is done for tests applied to each locus as well as to a multilocus estimate based on 10 loci.

In the island model, %RHo approaches α for relatively high migration rates (*i.e.*, $m = 10^{-1}$–$10^{-2} = 10$–100μ), in accordance with our *a priori* expectation that we should not detect a significant effect when $m \gg \mu$ (Figure 2A). On the contrary, for lower migration rates, mutation is no longer negligible compared to migration and the proportion of significant tests increases above α, reaching 88 and 100% when $m = 10^{-4}$ ($m = 0.1\mu$) for tests on a single locus or 10 loci, respectively (Figure 2A). Tests based on 10 loci seem actually quite powerful for typical sample sizes encountered in experimental studies (300 individuals here), as 100% of the tests were significant when $m = \mu$ and already 24% when $m = 10\mu$. Results of the two isolated population models are very similar to those of the island model if m is replaced by $1/t$ (Figure 2B). Here, however, tests seem less powerful than in the simulated island model (*e.g.*, for 10 loci, %RHo > 50% when $1/t \leq \mu$ in the isolated population model, and $m \leq 0.3\mu$ in the island model), which is likely due to the smaller sample size (100 *vs.* 300 individuals) and the lower number of populations sampled (2 *vs.* 10). BALLOUX and GOUDET (2002) showed indeed that the variance of R_{ST} increases substantially with fewer populations sampled. In the stepping-stone model, %RHo increases with the distance separating populations, but reaches a plateau beyond eight steps at ~60% for estimates based on 10 loci and only 20% for single-locus estimates (Figure 2C). Surprisingly, %RHo is already significantly larger than α for populations separated by just one step and exchanging migrants at a high rate ($m/2 = 0.05$) relative to the mutation rate ($\mu = 0.001$).

Usefulness of the test to determine the most appropriate statistics: To verify whether the test provides an adequate guideline to choose between R_{ST} and F_{ST} when assessing population differentiation, mean square errors (MSEs) of F_{ST} and R_{ST} were computed. The MSE is a synthetic measure of the efficiency of an estimator combining bias and variance (MSE = bias^2 + variance). It has already been used to compare the efficiency of F_{ST} and R_{ST} estimators (Balloux and Goudet 2002) or gene flow estimates based on F_{ST} or R_{ST} (Gaggiotti *et al.* 1999). MSEs were computed as $\Sigma(i - e)^2/n$, where i is the F_{ST} or R_{ST} estimate of the ith replicate, n is the number of replicates ($n = 200$), and e is the expected value given the demographic parameters. The expected value is $e = 1/(1 + 4Nmd/(d - 1))$ in the case of the island model (with $N = 100$ and $d = 10$), and $e = t/(2N + t)$ in the case of the isolated population model (with $N = 500$). These are the values expected for R_{ST} under SMM and for F_{ST} under IAM (or KAM) and a low mutation rate (Slatkin 1995; Rousset 1996). Note that e is not the expected F_{ST} under the conditions of the simulations (relatively high SMM and μ), but only a good approximation when mutation can be neglected.

For the island model and $\mu = 0.001$ (SMM), with migration rate varying from 0.0001 to 0.1, the ratio MSE(R_{ST})/MSE(F_{ST}) varied, respectively, from 0.06 to 2.1 for single-locus estimates and from 0.02 to 2.3 for multilocus estimates based on 10 loci. The migration rate at which MSE(R_{ST}) = MSE(F_{ST}) was between $m = 0.001$ and 0.002 for single-locus estimates and between $m = 0.003$ and 0.005 for multilocus estimates. As can be observed in Figure 2A, these migration rate limits under which R_{ST} performs better than F_{ST}, and above which the reverse occurs, closely match the migration rate under which the allele size permutation test becomes often significant (*i.e.*, %RHo $\geq$ 30%). The same pattern is observed for the isolated populations model: For t varying from 30 to 10,000 generations, MSE(R_{ST})/MSE(F_{ST}) varied from 2.37 to 0.41 and from 4.00 to 0.01 for single-locus and multilocus estimates, respectively, and MSE(R_{ST}) = MSE(F_{ST}) for $t = 2000$ (*i.e.*, $2/\mu$) and $t = 500$ (*i.e.*, $0.5/\mu$) for single-locus and multilocus estimates, respectively. Hence, the test becomes frequently significant when MSE(R_{ST}) is close to MSE(F_{ST}) (Figure 2B).

These results strongly suggest that the allele size permutation test is well suited to determine which of F_{ST} or R_{ST} is the most adequate for demographic parameters inferences, at least on the basis of the lowest MSE criterion. However, it must be pointed out that the statistic with lowest MSE is not necessarily the statistic that will provide the lowest MSE in the demographic estimate, because demographic estimates are usually not linear functions of F_{ST} or R_{ST}. For example, in the isolated population model, the $\tau = t/N$ estimates that can be derived using $\tau_F = 2F_{ST}/(1 - F_{ST})$ and $\tau_R = 2R_{ST}/(1 - R_{ST})$ give MSE(τ_R) > MSE(τ_F) for all simulated divergence time with single-locus estimates [τ_F can also be estimated as $-\ln(1 - F_{ST})$ (Reynolds *et al.* 1983), but this leads essentially to the same results]. This occurs because whenever F_{ST} or R_{ST} approaches 1, the inferred τ quickly takes enormous values, so that the impact of the larger variance of R_{ST} relative to F_{ST} is greatly amplified in the inferred τ, although τ_R is much less biased than τ_F for $\tau \geq 1$. The good news is that for multilocus estimates we obtained MSE(τ_R) = MSE(τ_F) for $t = 500$ and MSE(τ_R) < MSE(τ_F) for $t > 500$, as previously found for MSE(R_{ST}) = MSE(F_{ST}). Similarly, for the island model, where Nm can be estimated as $Nm_F = (1/F_{ST} - 1)/4$ and $Nm_R = (1/R_{ST} - 1)/4$, the m values corresponding to MSE(Nm_F) = MSE(Nm_R) were exactly equal to these obtained for MSE(R_{ST}) = MSE(F_{ST}) for both single- and multilocus estimates. Thus, the usefulness of the allele size permutation test to determine which of F_{ST} or R_{ST} is the most adequate for inferential purposes seems to be quite general, except probably with low sample size and/or low number of loci, when inferences are in any case doubtful because associated variances are too large.

Application examples: To illustrate the utility and power of the allele size permutation test with real data we present three examples of published data sets that we reanalyzed. These data were collected to assess population differentiation and check for isolation by distance in three different organisms. We computed global or pairwise F_{ST} and R_{ST} statistics as described above and applied the allele size permutation tests to obtain pR_{ST} values. These analyses were performed with SPAGeDi.

Biomphalaria pfeifferi, a selfing snail recently introduced in Madagascar: Biomphalaria pfeifferi, an intermediate host of a parasitic trematode causing intestinal bilharziasis, is a hermaphroditic freshwater snail distributed over most of Africa, the Middle East, and Madagascar. Madagascar was relatively recently invaded by this snail, probably as a result of human occupation a few hundred years ago (Charbonnel *et al.* 2002a). Moreover, according to a broad-scale survey of microsatellite variation throughout Madagascar, bottleneck (Cornuet and Luikart 1996) and admixture (Bertolle and Excoffier 1998) tests suggest that at least three independent introductions from genetically differentiated sources occurred (Charbonnel *et al.* 2002a). A small-scale study of microsatellite variation also reveals that populations experienced recurrent bottlenecks and that migration has been frequent within watersheds but rare among them (Charbonnel *et al.* 2002b). This population dynamic and the high selfing rate experienced by this snail explain the high genetic differentiation among populations observed in Madagascar: $F_{ST} = 0.80$ and 0.58 for broad and small scales, respectively (Charbonnel *et al.* 2002a,b).

In this particular context, we can formulate a hypothesis regarding the information content that microsatellite allele sizes could bear. Given the postulated recent introductions of this snail in Madagascar, we expect that mutation has not contributed to differentiation among populations originating from the same introduction but

TABLE 3

Differentiation among populations of *Biomphalaria pfeifferi* at different scales

Locus	No. of alleles	R_{ST}	pR_{ST} (95% C.I.)	F_{ST}
		Local scale		
Multilocus		0.571 NS	0.561 (0.438–0.676)	0.588
Bpf12	3	0.560	0.716 (0.457–0.930)	0.712
Bpf2	5	0.607 NS	0.594 (0.222–0.652)	0.645
Bpf1	6	0.483 NS	0.620 (0.417–0.852)	0.605
Bpf10	9	0.418 NS	0.578 (0.381–0.733)	0.596
Bpf9	14	0.589 NS	0.546 (0.380–0.714)	0.550
Bg16	16	0.453 NS	0.493 (0.234–0.675)	0.525
		Large scale		
Multilocus		0.960***	0.788 (0.676–0.903)	0.809
Bpf5	4	0.999	0.980 (0.929–0.999)	0.985
Bpf1	5	0.823 NS	0.798 (0.746–0.840)	0.798
Bpf2	8	0.954*	0.844 (0.637–0.960)	0.857
Bpf8	9	0.999***	0.835 (0.491–0.993)	0.895
Bg16	10	0.809 NS	0.775 (0.603–0.897)	0.783
BgE5	12	0.897***	0.718 (0.553–0.856)	0.724
Bpf10	13	0.823 NS	0.834 (0.546–0.971)	0.852
Bpf9	18	0.812**	0.624 (0.431–0.776)	0.636

The 95% confidence interval given with pR_{ST} is the 95% central pR_{ST} values obtained after random permutations of the allele sizes. P values of allele size permutation tests on R_{ST} are denoted as follows: NS, nonsignificant; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. No test was done for the loci with less than five alleles because the number of permutation configurations is too low to carry out a test at a 5% level.

has contributed to differentiation among populations originating from different introductions (at least if the source populations had diverged over enough time). The places and timing of the introductions are not known, but populations from a single watershed are likely to originate from a single introduction or, if genotypes from different introductions mixed in a watershed, migration within the watershed is likely to have prevented the buildup of a phylogeographical pattern at this scale. Therefore, we can expect R_{ST} to be close to F_{ST} for populations belonging to the same watershed and significantly larger than F_{ST} for populations from different watersheds when the latter were originally colonized by individuals from independent introductions.

To test this hypothesis, we reanalyzed data from small-scale and large-scale studies by Charbonnel *et al.* (2002a,b). Global R_{ST} and F_{ST} values as well as pairwise R_{ST} and F_{ST} values between populations were computed. Distinguishing pairs of populations within or among watersheds, pairwise values were regressed on spatial distances (Mantel tests were used to assess the significance of the regression slopes), and average pairwise values were computed for a set of distance classes (defined in such a way that each contained ~33 pairs of populations). One thousand random permutations of the allele sizes provided a distribution of pR_{ST} values, 95% confidence intervals covering the 25th to the 975th ordered values, and P values testing if $R_{ST} > pR_{ST}$.

Multilocus R_{ST} values are significantly higher than mean pR_{ST} at a broad scale but not at a local scale (Table 3). Applied to each locus, these tests were also significant for four out of eight loci at the broad scale but for none at the local scale.

The analysis of average pairwise multilocus F_{ST} and R_{ST} values per distance class at the broad scale shows the following (Figure 3):

1. Differentiation between populations occupying the same watershed is much lower than that between populations from different watersheds, even for populations separated by the same spatial distance. This is in line with the higher migration rate detected within watersheds than among them (Charbonnel 2002b).
2. A pattern of isolation by distance is detected within watersheds for both F_{ST} and R_{ST} (Mantel tests: $P = 0.007$ and 0.021, respectively). Among watersheds, such a pattern is not detected for F_{ST} but is for R_{ST} (Mantel tests: $P = 0.18$ and 0.002, respectively).
3. Within watersheds, R_{ST}'s are not significantly higher than pR_{ST}'s, whereas among watersheds, R_{ST}'s are significantly higher than pR_{ST}'s for all distance classes but the first one.
4. Average pairwise pR_{ST} values are always somewhat lower than pairwise F_{ST} values but they follow closely their pattern of variation with spatial distance.

In conclusion, at a local scale, R_{ST} values are close to F_{ST} values, and allele size permutation tests do not reveal any significant contribution of stepwise mutations to population differentiation. On the contrary, at a large

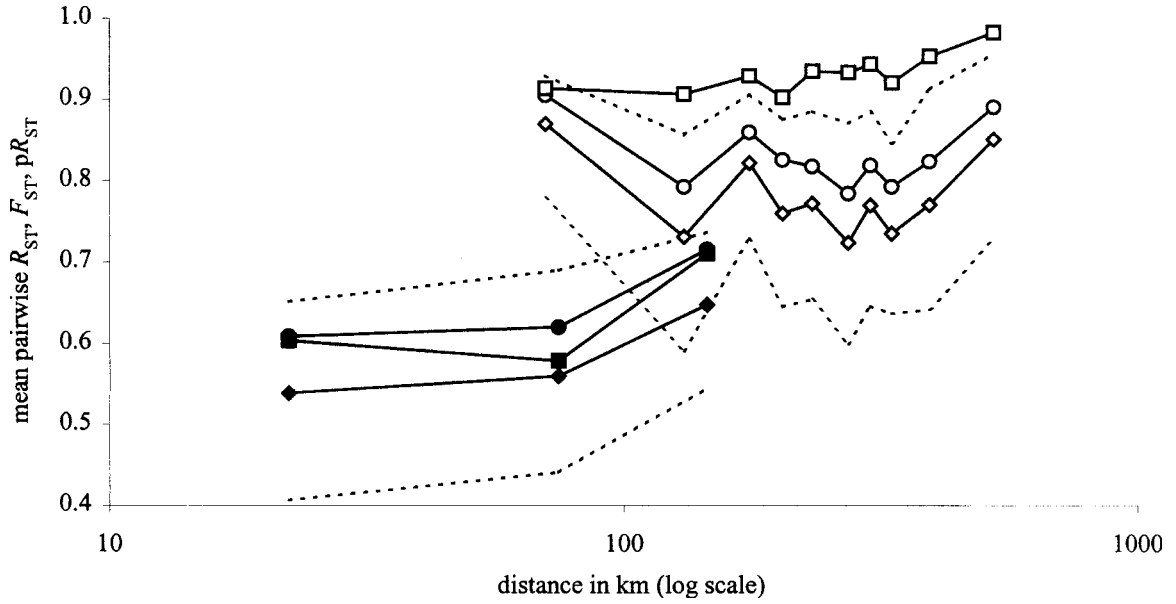

FIGURE 3.—Average pairwise F_{ST} (○ and ●), R_{ST} (□ and ■), and mean pR_{ST} (◇ and ◆) values among populations of *Biomphalaria pfeifferi* throughout Madagascar for a set of distance classes, distinguishing comparisons between populations within watersheds (●, ■, ◆) and among watersheds (○, □, ◇). The dotted lines represent the range of the 95% central ordered pR_{ST} values (*i.e.*, after allele size randomization). Each distance class contains 32–35 pairs of populations.

scale, R_{ST} values are substantially higher than F_{ST} values and allele size permutation tests demonstrate that shifts in average allele sizes contribute significantly to population differentiation. Significant tests on R_{ST} values are expected if populations had diverged for a sufficiently long time and/or if populations exchanged migrants at a rate similar or inferior to the mutation rate. The results are thus very consistent with *a priori* expectations given that (1) at a large scale, both these conditions are probably met because populations far apart in Madagascar probably originated from relatively recent and independent introductions from source continental populations isolated for a long time, and migration rate is low among watersheds, and (2) at a local scale, particularly within watersheds, none of these conditions are likely to be met.

Fraxinus excelsior, a widespread European tree: Fraxinus excelsior (Oleaceae, common ash) is a widespread European wind-pollinated tree species found mostly in floodplain locations and with a scattered distribution within natural forests. The distribution of chloroplastic DNA (cpDNA) haplotypes throughout Europe suggests that *F. excelsior* was located in at least three different refuges during the last ice age, one putative refuge being the Balkan area (G. G. VENDRAMIN, unpublished data). HEUERTZ *et al.* (2001) analyzed microsatellite polymorphism in 10 Bulgarian populations (Balkan area) from three regions (321 individuals). Populations were separated by 0.5–22 km within regions and 120–300 km among regions.

In the absence of evidence of long-term divergence between Bulgarian populations (no evidence of different refuges), and given that gene flow should be relatively extended in a wind-pollinated species, we may expect that stepwise-like mutations have not contributed significantly to population differentiation in Bulgaria. The data set of HEUERTZ *et al.* (2001) was thus reanalyzed to compare average pairwise F_{ST} and R_{ST} values between populations, distinguishing pairs within and among Bulgarian regions, and testing R_{ST} values by allele size permutations (1000 randomizations).

Mean pairwise multilocus estimates were equal to $F_{ST} = 0.074$, $R_{ST} = 0.091$ within regions and $F_{ST} = 0.097$, $R_{ST} = 0.180$ among regions (Figure 4). Hence, whereas differentiation increases slightly from small to large geographical scales according to F_{ST}, it nearly doubles according to R_{ST}. Moreover, average pairwise R_{ST} is much larger than F_{ST} among regions, but only slightly larger than F_{ST} within regions. Within regions, observed R_{ST}'s are always within the 95% range of central pR_{ST}, but among regions, the multilocus R_{ST} estimate as well as the estimate for locus FEM19 is larger than the 95% range of pR_{ST} (Figure 4), demonstrating that stepwise-like mutations contributed to population differentiation at the large geographical scale for at least one locus.

Several causes may account for the significant allele size effect on population differentiation among regions in Bulgaria, for example:

1. The pattern may reflect isolation by distance. However, it seems unlikely that migration rate among regions is weak compared to the mutation rate given that pollen is wind dispersed.
2. The pattern may be due to postglacial recolonization from different refuges. There is, however, no evi-

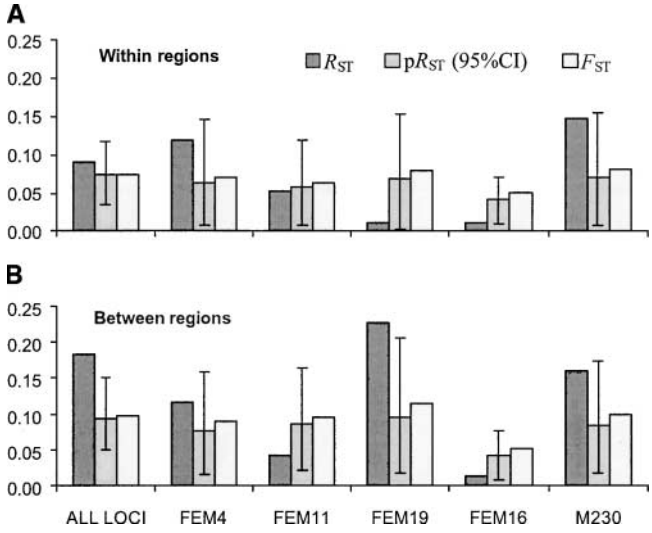

FIGURE 4.—Mean pairwise R_{ST}, mean pR_{ST}, and F_{ST} values between Bulgarian populations of *Fraxinus excelsior* for populations belonging to the same region (A) or different regions (B). Values are given for each locus and the multilocus estimates. Bars of pR_{ST} indicate the mean pR_{ST} values over 1000 allele size permutations, and the corresponding intervals give the range of the 95% central pR_{ST} values.

dence of different refuges from the maternally inherited cytoplasmic DNA as the same unique haplotype occurs in all three regions (M. HEUERTZ, unpublished data).

3. The pattern may reflect human-mediated introduction of Fraxinus from remote regions.
4. The pattern may reflect locally occurring hybridization between *F. excelsior* and a related species such as *F. angustifolia* or *F. pallisiae*. Given that a total of four ash species (the former three and *F. ornus*) are found in Bulgaria and that different species occur in the same forests (M. HEUERTZ, personal observation), this latter hypothesis merits further investigation. In any case, the observation that a significant effect of stepwise-like mutations is observed on a large scale but not on a small one remains very consistent with *a priori* expectations, as nearby populations should exchange genes at a relatively high rate.

Centaurea corymbosa, a rare and narrow-ranged cliff-dwelling herb: Centaurea corymbosa (Asteraceae) is a short-lived perennial herb species distributed over a very narrow range (within a 3-km^2 area of a calcareous massif along the French Mediterranean coast), where it occurs in only six small populations (COLAS *et al.* 1997). It has specialized into an extreme habitat: the top of limestone cliffs where few other plant species survive. On more fertile ground, *C. corymbosa* is outcompeted, so that suitable habitat is highly fragmented, appearing as small islands dispersed in the landscape. Given that the species occupies only a small fraction of these "islands" (the whole massif extends over 50 km^2), colonization ability must be very limited, probably as a consequence of limited seed dispersal ability and the self-incompatibility system that prevents a potential newcomer from founding a new population on its own (COLAS *et al.* 1997; FRÉVILLE *et al.* 2001). Patterns of isozyme (COLAS *et al.* 1997) and microsatellite (FRÉVILLE *et al.* 2001) variation show high levels of differentiation among populations, with $F_{ST} = 0.35$ and 0.23, respectively, despite the narrow range of the species (2.3 km between the two most distant populations). High differentiation at such a small scale cannot be attributed to the mating system as the species is self-incompatible. It most likely results from small population sizes and low gene flow among populations. It might also be a consequence of more or less recurrent bottlenecks when new populations are founded (although the turnover should be relatively slow, given that no population extinction or foundation has been observed since 1994, when *C. corymbosa* populations began to be closely surveyed, and herbarium data show that five of the six populations were known >100 years ago).

In this context it is interesting to question whether gene flow among populations is sufficiently low to permit divergence by mutations. The higher observed F_{ST} value at allozyme loci than at microsatellite loci could indeed be caused by high mutation rates of microsatellites, provided that $\mu \geq m$. FRÉVILLE *et al.* (2001) pointed out that this hypothesis was also supported by the fact that F_{ST} values at the two most polymorphic microsatellite loci (12B1 and 21D9, Table 4), the ones likely to have the highest mutation rates, were lower than those for the two loci with intermediate levels of polymorphism (13D10 and 28A7, Table 4).

The allele size randomization procedure is adequate to address this question. Therefore, global R_{ST}, pR_{ST}, and F_{ST} were computed for microsatellite loci as described above, and R_{ST} was compared against the distribution of 1000 pR_{ST} values. Permutation tests did not detect any R_{ST} value significantly $> pR_{ST}$ (Table 4). This suggests thus that differentiation is caused mainly by drift and that gene flow, m, and/or the reciprocal of divergence time, $1/t$, are large compared to the mutation rate, μ. This result also implies that F_{ST} should be a better estimator than R_{ST} of population differentiation for this species. Actually, given the small population sizes (COLAS *et al.* 1997, 2001), drift is expected to be high. For example, if populations had effective sizes of $\sim$100 individuals (there is actually much variance among populations) and conformed to an island model (there are actually some isolation-by-distance effects), a value of $m = 0.006$ would account for the observed F_{ST}, a value larger than typical microsatellite mutation rates (10^{-3}–10^{-4}). Assuming that these populations have been in place for a sufficiently long time to potentially permit differentiation by mutations (shifting allele sizes), the absence of such mutation-driven differentiation also suggests that the migration rate is larger than the muta-

TABLE 4

Differentiation among populations of *Centaurea corymbosa*, estimated by global R_{ST}, mean pR_{ST}, and F_{ST} values per locus and for a multilocus average

Locus	No. of alleles	R_{ST}	pR_{ST}	F_{ST}
Multilocus		0.259 NS	0.222 (0.119–0.342)	0.228
17E3	3	0.124	0.133 (0.124–0.153)	0.130
13B7	3	0.096	0.082 (0.048–0.096)	0.094
13D10	5	0.273 NS	0.339 (0.177–0.587)	0.341
28A7	7	0.288 NS	0.261 (0.066–0.526)	0.272
21D9	12	0.230 NS	0.182 (0.029–0.392)	0.181
12B1	15	0.276 NS	0.194 (0.020–0.399)	0.194

The 95% confidence interval given with pR_{ST} is the 95% central pR_{ST} values obtained after random permutations of the allele sizes. No test was done for loci with three alleles because the number of permutation configurations is too low to carry out a test at a 5% level. *P* values of allele size permutation tests on R_{ST} are denoted as follows: NS, nonsignificant ($P > 0.05$).

tion rate, so that new mutation variants spread over all populations.

Nonsignificant tests could also be due to a lack of power, so the test should be applied to additional microsatellite loci to confirm these results (presently, only four out of six loci had a sufficient number of alleles to carry out permutation tests). Deviation from a SMM at some loci could also reduce the power of the test. For example, the dinucleotide locus 28A7 has six alleles with sizes following a sequence of one repeat step plus one allele at least six repeats smaller than the other ones. Although this pattern is not necessarily incompatible with a pure SMM (*e.g.*, Donnelly 1999), it might suggest that a mutation of large effect created the outsider allele.

DISCUSSION

Comparison between measures of differentiation: Comparisons of F_{ST} with R_{ST} values on microsatellite data have already been suggested for checking the importance of mutation *vs.* migration rates (*e.g.*, Michalakis and Veuille 1996; Ross *et al.* 1997; Estoup *et al.* 1998). For example, in the brown trout (*Salmo trutta*), populations sampled at a microgeographic scale showed similar R_{ST} and F_{ST} estimates, whereas populations sampled at a macrogeographic scale showed significantly higher R_{ST} compared to F_{ST}, indicating that mutation becomes important relative to migration at this scale (Estoup and Angers 1998). Similarly, in a review of F_{ST}-R_{ST} data analyses, Lugon-Moulin *et al.* (1999) showed that R_{ST} and F_{ST} are generally similar when the level of differentiation is low, whereas R_{ST} is often superior to F_{ST} when differentiation is high. The same trend was observed in two of the data sets reanalyzed in the present article (*F. excelsior* and *B. pfeifferi*).

To compare multilocus F_{ST} and R_{ST} estimates, Estoup and Angers (1998) applied a nonparametric Wilcoxon signed-rank test on single-locus F_{ST} and R_{ST} estimates, and Lugon-Moulin *et al.* (1999) used a bootstrapping procedure over loci. These approaches assume that F_{ST} should be equal to R_{ST} if mutations can be neglected, which is true for the corresponding parameters (Rousset 1996), but not necessarily true for the estimators because they can be subject to different bias. Actually, a difference in bias was detected in the simulation results where F_{ST} and R_{ST} were computed for two independent samples from a single population (*i.e.*, no actual differentiation): The percentages of loci (>200) with $R_{ST} < F_{ST}$ were equal to 65 and 69% under KAM and SMM, respectively, resulting in significant sign tests, although, as parameters, F_{ST} and R_{ST} were both equal to zero. The allele size permutation test has the advantages that (1) a test can be applied to each locus (mutation rate and process are locus specific) and (2) R_{ST} is compared to the same statistic but computed on data with randomized allele sizes, so that potential statistical bias on the compared statistics should be identical.

Comparison between F_{ST} and R_{ST} is similar to comparing G_{ST} with N_{ST} on haplotypes (*i.e.*, DNA sequences or other nonrecombinant DNA variants, such as mitochondrial or chloroplastic DNA; Pons and Petit 1996). Indeed, G_{ST} is a measure of differentiation (very similar to F_{ST}) between haplotypes using "unordered" alleles (*i.e.*, not accounting for the similarities between haplotypes) whereas N_{ST} is a measure based on "ordered" alleles (*i.e.*, accounting for the similarities between haplotypes). Mathematically, $G_{ST} \equiv (h_T - h_w)/h_T$ and $N_{ST} \equiv (v_T - v_w)/v_T$, where h and v are measures of genetic diversity and subscripts $_T$ and $_w$ refer to diversity measured over the total set of populations and within population, respectively (see Pons and Petit 1996 for details and parameters estimation). The diversity measures h (heterozygosity) depend only on haplotype frequencies and are of the form $h = 1 - \Sigma p_i^2$, where p_i is the ith allele frequency, which is equivalent to $h = \Sigma_i \Sigma_j \pi_{ij} p_i p_j$,

where $\pi_{ij} = 0$ if $i = j$ and $\pi_{ij} = 1$ otherwise. The diversity measures v depend also on haplotype divergence and are of the form $v = \Sigma_i \Sigma_j \pi_{ij} p_i p_j$, where π_{ij} now represents a degree of divergence between haplotypes i and j ($\pi_{ij} = 0$ if $i = j$ but otherwise π_{ij} varies, being, for example, proportional to the number of site differences between i and j). N_{ST} is expected to be $> G_{ST}$ when similar haplotypes (*i.e.*, haplotype pairs with low π_{ij}) are associated geographically; otherwise they should have identical expectations. Thus, when comparing R_{ST} with F_{ST} or N_{ST} with G_{ST}, measures of differentiation based on ordered *vs.* unordered alleles are compared, and the importance of mutation relative to other causes of genetic differentiation (*i.e.*, gene flow and divergence time) can be assessed. Pons and Petit (1996) proposed a parametric test to compare G_{ST} and N_{ST}, but a nonparametric test based on random permutations of genetic distances between haplotypes was later used and proved to be more efficient (Burban *et al.* 1999; Petit *et al.* 2002). The allele size permutation test proposed in this article is actually identical to permuting genetic distances between alleles.

Impact of deviations from a pure SMM on the power of the test: In all the simulations realized to assess the power of the test, a strict SMM was considered. However, the microsatellite mutation process is known to deviate from a strict SMM (Lehman *et al.* 1996; Wierdl *et al.* 1997; Zhivotovsky *et al.* 1997; Estoup and Angers 1998). For example, the polymorphism at dinucleotide microsatellite loci across the human genome is not consistent with a strict SMM but fits a model composed of a majority of single-step mutations and a small proportion of multistep mutations (Renwick *et al.* 2001). Similarly, allele size constraints were invoked to explain the polymorphism at human trinucleotide loci (Deka *et al.* 1999). One advantage of the allele size permutation test is that it remains valid under these deviations, the only requirement being that mutation favors short allele size changes when testing for the impact of mutation relative to drift (Table 2, case 2). Nevertheless, the power of the test would likely be reduced if the mutation process contained a significant proportion of mutations of large effect or if the range of allele sizes was constrained. To assess the loss of power of the test under these conditions, additional simulations of the island model were run allowing (1) for constraints on the allele size range (range = 30, 8, or 6) and (2) for nonstepwise mutations in the form of a proportion (20%) of double-step mutations (DSMs) or random mutations (KAM-like). Results for $m = 0.001$ and $\mu = 0.001$ are given in Table 5. Under these parameters, the range of allele sizes under SMM and without constraint varies between 5 and 14 per locus, with an average close to 8. Hence, adding a range constraint of 30, 8, and 6 can be interpreted as no, moderate, and strong range constraints, respectively. As expected, deviations from SMM resulted in a reduction of the power of the test (Table 5). However, the reduction was substantial only under the strong range constraint or when KAM-like mutations were included. In the latter case, the effect was more pronounced when the allele size range was unconstrained, a condition in which KAM-like mutations cause larger allele size changes. Hence, these results suggest that the allele size permutation test remains quite powerful under allele size constraint and multistep mutations. Deviations from the SMM are probably a more important concern when inferring demographic parameters. Indeed, if a significant test means that an F_{ST}-based demographic inference is likely to be biased, it does not demonstrate that an R_{ST}-based inference will be less biased, because the relationships used in R_{ST}-based inferences usually assume a strict SMM or GSM (see also Estoup *et al.* 2002 for the impact of the SMM and its deviations on size homoplasy).

Impact of nonequilibrium dynamics and selection: In the simulations performed, constant population size and/or mutation-drift equilibrium were assumed. These assumptions are also made when inferring demographic parameters (m or t) from the statistics measuring genetic differentiation or genetic distances. In many natural populations, these assumptions are not satisfied, potentially leading to strongly biased estimates (*e.g.*, Whitlock and McCauley 1999; Zhivotovsky 2001). However, because it does not rely on such assumptions, the allele size permutation test is expected to remain exact with respect to these violations in the sense that, whatever the demographic processes, the test will indicate whether stepwise mutations contributed significantly to genetic differentiation. It is, however, possible that the relative magnitude of μ with respect to m or $1/t$ at which the test becomes significant is affected by fluctuations of demographic parameters. This problem merits further investigations.

Neutrality of genetic markers with respect to natural selection was also assumed throughout this article. However, there are some lines of evidence that certain microsatellite markers are involved in functional roles and could therefore be subject to natural selection (*e.g.*, Kashi and Soller 1999). If selection acts on a microsatellite locus, it could have a major impact on the outcome of the allele size randomization test as soon as it selects for different allele size ranges in different populations, causing the test to be significant even if mutation-mediated differentiation is negligible relative to drift. On the contrary, if selection selects for the same range of allele sizes everywhere, it will essentially cause constraints on the range of allele sizes. As shown previously, the test is fairly robust to such constraints.

Other applications of the allele size permutation test: We suggested previously that the test can also be useful in choosing between statistics used for phylogenetic inference. For example, Ds (Nei 1972), based on allele identity information, and $(\delta\mu)^2$ (Goldstein *et al.*

TABLE 5

Impact of deviations from the stepwise mutation model (SMM) on the power of the allele size permutation test

Mutation model	Allele size range	1 locus			10 loci		
		F_{ST}	R_{ST}	%RHo	F_{ST}	R_{ST}	%RHo
100% SMM	30	0.55	0.67	44	0.55	0.68	96
		(0.08)	(0.14)		(0.02)	(0.04)	
80% SMM	30	0.56	0.69	37	0.57	0.72	98
20% DSM		(0.09)	(0.14)		(0.03)	(0.05)	
80% SMM	30	0.54	0.59	20	0.55	0.62	63
20% KAM		(0.08)	(0.16)		(0.02)	(0.05)	
100% SMM	8	0.55	0.68	36	0.55	0.70	98
		(0.09)	(0.14)		(0.03)	(0.04)	
80% SMM	8	0.57	0.66	29	0.57	0.68	91
20% DSM		(0.10)	(0.15)		(0.03)	(0.05)	
80% SMM	8	0.58	0.65	26	0.58	0.66	77
20% KAM		(0.09)	(0.15)		(0.03)	(0.05)	
100% SMM	6	0.55	0.65	23	0.56	0.67	86
		(0.10)	(0.15)		(0.03)	(0.05)	

Deviations occur in the form of (1) constraints on allele size range and (2) occurrence (at a 20% rate) of double-step mutations (DSM) or random mutations to any of the possible alleles (KAM). For single-locus and multilocus (10 loci) estimates, averages (standard deviations) of F_{ST} and R_{ST} values are reported, as well as the percentage of significant tests (%RHo) at a 5% confidence level (over 200 replicates). Simulations correspond to an island model (see text for details) with migration rate $m = 0.001$ (mutation rate $\mu = 0.001$).

1995a,b; Goldstein and Pollock 1997), based on allele size information, are genetic distances between populations with expectation $2\mu t$, but under the IAM for *Ds* and under the SMM for $(\delta\mu)^2$. In the case of microsatellites undergoing SMM-like mutations, *Ds* is strongly biased for large t (Goldstein *et al.* 1995b), but for small t it may remain relatively little biased and has a lower variance than $(\mu\delta)^2$ (Takezaki and Nei 1996). Could the allele size permutation test applied to R_{ST} be useful for choosing between *Ds* and $(\mu\delta)^2$? Using our simulation results of the isolated populations model (results not shown), the analysis of the MSE of divergence time estimates based on *Ds vs.* $(\mu\delta)^2$ permits us to conclude the following: A nonsignificant test suggests that *Ds* should be preferred for its low bias and variance. A significant test suggests that *Ds* is biased whereas $(\mu\delta)^2$ is essentially unbiased but, in terms of MSE, *Ds* still performs better unless t is very large, especially with a low number of loci. Hence, for the purpose of choosing between *Ds* and $(\mu\delta)^2$, the test is truly useful only when it gives a nonsignificant result (see also Takezaki and Nei 1996).

Assessing the significance of stepwise-like mutations to genetic differentiation may also have applications when studying inbreeding depression. The latter is often investigated by measuring the correlation between individual fitness and some measure of inbreeding: either the observed heterozygosity, H, or the average squared difference in repeat numbers between alleles within individuals, d^2 (Goudet and Keller 2002). Tsitrone *et al.* (2001) demonstrated that H should perform better than d^2 in most realistic conditions, except when individuals result from the recent admixture of populations having differentiated for a long time (with $N\mu \gg 1$, where N is the population size before admixture). Potentially, the allele permutation test might help identify such situations where d^2 performs better than H. If the source populations are known, it could be applied to an R_{ST} estimate between these populations. Otherwise, it could be applied to an R_{IS} estimate (the correlation of allele sizes between genes sampled within individuals) for the population after admixture. Although it is not obvious that a significant test would necessarily indicate that d^2 performs better than H, a nonsignificant test indicates that allele size is uninformative and, hence, H should surely perform better than d^2.

Beyond its practical use in choosing among statistics, the test can provide insights into the evolutionary interpretation of data sets by giving information on the relative values of the mutation rate compared to the migration rate or the time since population divergence. Simulations showed indeed that the test becomes quite powerful when the mutation rate, μ, is higher than the migration rate, m, or the reverse of the number of generations since population divergence, $1/t$. This is useful information, especially if mutation rates are roughly known, because gene flow estimates directly derived from F_{ST} or R_{ST} estimates are always expressed in terms of Nm products, where N, the effective population size, is often difficult to assess. However, only qualitative

insights on m or $1/t$ can probably be extracted from the test, because the exact ratio μ over m or μ over $1/t$ at which the test becomes highly powerful depends on the sample size, the number of loci, and probably the diversity of each locus. With many loci, a value of $\mu = 0.1m$ can already lead to a significant test, whereas with one locus and a small sample size, μ might exceed $10m$ to obtain a significant test with high probability.

We thank A. Estoup, J. Goudet, P. Jarne, and X. Vekemans, as well as two anonymous reviewers, for helpful comments and suggestions on this manuscript. O. Hardy is a postdoctoral researcher from the Belgian National Fund for Scientific Research. H. Fréville was supported through a European Community Marie Curie Fellowship.

LITERATURE CITED

Balloux, F., 2001 EASYPOP (version 1.7). A computer program for the simulation of population genetics. J. Hered. **92:** 301–302 (http://www.unil.ch/izea/softwares/easypop.html).

Balloux, F., and J. Goudet, 2002 Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. Mol. Ecol. **11:** 771–783.

Balloux, F., and N. Lugon-Moulin, 2002 The estimation of population differentiation with microsatellite markers. Mol. Ecol. **11:** 155–165.

Bertolle, G., and L. Excoffier, 1998 Inferring admixture proportions from molecular data. Mol. Biol. Evol. **15:** 1298–1311.

Burban, C., R. J. Petit, E. Carcreff and H. Jactel, 1999 Rangewise variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera: Matsucoccidae) in relation to the genetic structure of its host. Mol. Ecol. **8:** 1593–1602.

Charbonnel, N., B. Angers, R. Rasatavonjizay, P. Brémond, C. DeBain *et al.*, 2002a The influence of mating system, demography, parasites and colonization on the population structure of Biomphalaria pfeifferi in Madagascar. Mol. Ecol. **11:** 2213–2228.

Charbonnel, N., B. Angers, R. Rasatavonjizay, P. Brémond and P. Jarne, 2002b Evolutionary aspects of the metapopulation dynamics of *Biomphalaria pfeifferi*, the intermediate host of *Schistosoma mansoni*, in Madagascar. J. Evol. Biol. **15:** 248–261.

Colas, B., I. Olivieri and M. Riba, 1997 *Centaurea corymbosa*, a cliff-dwelling species tottering on the brink of extinction: a demographic and genetic study. Proc. Natl. Acad. Sci. USA **94:** 3471–3476.

Colas, B., I. Olivieri and M. Riba, 2001 Spatio-temporal variation of reproductive success and conservation of the narrow-endemic *Centaurea corymbosa* (Asteraceae). Biol. Conserv. **99:** 375–386.

Cornuet, J.-M., and G. Luikart, 1996 Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics **144:** 2001–2014.

Crow, J. F., and K. Aoki, 1984 Group selection for a polygenic behavioural trait: estimating the degree of population subdivision. Proc. Natl. Acad. Sci. USA **81:** 6073–6077.

Deka, R., S. Guangyun, D. Smelser, Y. Zhong and M. Kimmel, 1999 Rate and directionality of mutations and effects of allele size constraints at anonymous, gene-associated, and disease-causing trinucleotide loci. Mol. Biol. Evol. **16:** 1166–1177.

Donnelly, P., 1999 The coalescent and microsatellite variability, pp. 116–128 in *Microsatellites: Evolution and Applications*, edited by D. B. Goldstein and C. Schlötterer. Oxford University Press, Oxford.

Ellegren, H., 2000 Heterogeneous mutation processes in human microsatellite DNA sequences. Nat. Genet. **24:** 400–402.

Estoup, A., and B. Angers, 1998 Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations, pp. 55–86 in *Advances in Molecular Ecology*, edited by G. R. Carvalho. IOS Press, Amsterdam.

Estoup, A., and J.-M. Cornuet, 1999 Microsatellite evolution: inferences from population data, pp. 49–65 in *Microsatellites: Evolution and Applications*, edited by D. B. Goldstein and C. Schlötterer. Oxford University Press, Oxford.

Estoup, A., F. Rousset, Y. Michalakis, J.-M. Cornuet, M. Adriamanga *et al.*, 1998 Comparative analysis of microsatellite and allozyme markers: a case study investigating microgeographic differentiation in brown trout (*Salmo trutta*). Mol. Ecol. **7:** 339–353.

Estoup, A., P. Jarne and J.-M. Cornuet, 2002 Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol. Ecol. **11:** 1591–1604.

Excoffier, L., 2001 Analysis of population subdivision, pp. 271–324 in *Handbook of Statistical Genetics*, edited by D. J. Balding, M. Bishop and C. Cannings. John Wiley & Sons, Chichester, UK.

Feldman, M. W., J. Kumm and J. Pritchard, 1999 Mutation and migration in models of microsatellite evolution, pp. 98–115 in *Microsatellites: Evolution and Applications*, edited by D. B. Goldstein and C. Schlötterer. Oxford University Press, Oxford.

Fréville, H., F. Justy and I. Olivieri, 2001 Comparative allozyme and microsatellite population structure in a narrow endemic plant species, *Centaurea corymbosa* Pourret (Asteraceae). Mol. Ecol. **10:** 879–889.

Gaggiotti, O. E., O. Lange, K. Rassmann and C. Gliddon, 1999 A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. Mol. Ecol. **8:** 1513–1520.

Goldstein, D. B., and D. D. Pollock, 1997 Launching microsatellites: a review of mutation processes and method for phylogenetic inference. J. Hered. **88:** 335–342.

Goldstein, D. B., A. R. Linares, M. W. Feldman and L. L. Cavalli-Sforza, 1995a Genetic absolute dating based on microsatellites and the origin of modern humans. Proc. Natl. Acad. Sci. USA **92:** 6723–6727.

Goldstein, D. B., A. R. Linares, M. W. Feldman and L. L. Cavalli-Sforza, 1995b An evaluation of genetic distances for use with microsatellite loci. Genetics **139:** 463–471.

Goudet, J., and L. Keller, 2002 The correlation between inbreeding and fitness: Does allele size matter? Trends Ecol. Evol. **17:** 201–202.

Hardy, O. J., and X. Vekemans, 2002 SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol. Ecol. Notes **2:** 618–620.

Heuertz, M., J.-F. Hausman, I. Tsvetkov, N. Frascaria-Lacoste and X. Vekemans, 2001 Assessment of genetic structure within and among Bulgarian populations of the common ash (*Fraxinus excelsior* L.). Mol. Ecol. **10:** 1615–1623.

Jarne, P., and J. L. Lagoda, 1996 Microsatellites, from molecules to populations and back. Trends Ecol. Evol. **11:** 424–429.

Kashi, Y., and M. Soller, 1999 Functional roles of microsatellites and minisatellites, pp. 10–23 in *Microsatellites: Evolution and Applications*, edited by D. B. Goldstein and C. Schlötterer. Oxford University Press, Oxford.

Kimmel, M., and R. Chakraborty, 1996 Measures of variation at DNA repeat loci under a general stepwise mutation model. Theor. Popul. Biol. **50:** 345–367.

Kimmel, M., R. Chakraborty, D. N. Stivers and R. Deka, 1996 Dynamics of repeat polymorphisms under a forward-backward mutation model: within- and between-population variability at microsatellite loci. Genetics **143:** 549–555.

Lehman, T., W. A. Hawlay and F. H. Collins, 1996 An evaluation of evolutionary constraints on microsatellite loci using null alleles. Genetics **144:** 1155–1163.

Lugon-Moulin, N., H. Brünner, A. Wyttenbach, J. Hausser and J. Goudet, 1999 Hierarchical analyses of genetic differentiation in a hybrid zone of *Sorex araneus* (Instectivora: Soricidae). Mol. Ecol. **8:** 419–431.

Michalakis, Y., and L. Excoffier, 1996 A genetic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics **142:** 1061–1064.

Michalakis, Y., and M. Veuille, 1996 Length variation of CAG/CAA trinucleotide repeats in natural populations of *Drosophila melanogaster* and its relation to recombination rate. Genetics **143:** 1713–1725.

Nei, M., 1972 Genetic distance between populations. Am. Nat. **106:** 283–292.

Petit, R. J., U. M. Csaikl, S. Bordács, K. Burg, E. Coart *et al.*, 2002 Chloroplast DNA variation in European white oaks. Phylogeography and patterns of diversity based on data from over 2600 populations. For. Ecol. Manage. **156:** 5–26.

Pons, O., and R. J. Petit, 1996 Measuring and testing genetic differentiation with ordered *versus* unordered alleles. Genetics **144:** 1237–1245.

Renwick, A., L. Davison, H. Spratt, J. P. King and M. Kimmel, 2001 DNA dinucleotide evolution in humans: fitting theory to facts. Genetics **159:** 737–747.

Reynolds, J., B. S. Weir and C. C. Cockerham, 1983 Estimation of the coancestry coefficient: basis for a short-term genetic distance. Genetics **105:** 767–779.

Ross, K. G., M. J. Krieger, D. D. Shoemaker, E. L. Vargo and L. Keller, 1997 Hierarchical analysis of genetic structure in native fire ant populations: results from three classes of molecular markers. Genetics **147:** 643–655.

Rousset, F., 1996 Equilibrium values of measures of population subdivision for stepwise mutation processes. Genetics **142:** 1357–1362.

Rousset, F., 1997 Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. Genetics **145:** 1219–1228.

Shriver, M. D., M. W. Smith, L. Jin, A. Marcini, J. M. Akey *et al.*, 1995 A novel measure of genetic distance for highly polymorphic tandem repeat loci. Mol. Biol. Evol. **12:** 914–920.

Slatkin, M., 1985 Gene flow in natural populations. Annu. Rev. Ecol. Syst. **16:** 393–430.

Slatkin, M., 1995 A measure of population subdivision based on microsatellite allele frequencies. Genetics **139:** 457–462.

Takezaki, N., and M. Nei, 1996 Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics **144:** 389–399.

Tsitrone, A., F. Rousset and P. David, 2001 Heterosis, marker mutational processes and population inbreeding history. Genetics **159:** 1845–1859.

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

Whitlock, M. C., and D. E. McCauley, 1999 Indirect measures of gene flow and migration: Fst≠1/(4Nm+1). Heredity **82:** 117–125.

Wierdl, M., M. Dominska and T. D. Petes, 1997 Microsatellite instability in yeast: dependence on the length of the microsatellite. Genetics **146:** 769–779.

Wright, S., 1965 The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution **19:** 395–420.

Xu, X., M. Peng, Z. Fang and X. P. Xu, 2000 The direction of microsatellite mutations is dependent upon allele length. Nat. Genet. **24:** 396–399.

Zhivotovsky, L. A., 2001 Estimating divergence time with the use of microsatellite genetic distances: impacts of population growth and gene flow. Mol. Biol. Evol. **18:** 700–709.

Zhivotovsky, L. A., and M. W. Feldman, 1995 Microsatellite variability and genetic distances. Proc. Natl. Acad. Sci. USA **92:** 11549–11552.

Zhivotovsky, L. A., M. W. Feldman and S. A. Grishechkin, 1997 Biased mutations and microsatellite variation. Mol. Biol. Evol. **14:** 926–933.

Zhu, Y., J. E. Strassmann and D. C. Queller, 2000 Insertions, substitutions, and the origin of microsatellites. Genet. Res. **76:** 227–236.

Communicating editor: M. W. Feldman