Statistical properties of population differentiation estimators under stepwise mutation in a finite island model

F. BALLOUX*‡ and J. GOUDET†

**University of Bern, CH-3032 Hinterkappelen-Bern, Switzerland, †Institute of Ecology, Biology Building, University of Lausanne, CH-1015 Lausanne, Switzerland, ‡University of Edinburgh, Institute of Cell, Animal and Population Biology, King's Buildings, West Mains Road, Edinburgh EH9 3JT, UK*

Abstract

Microsatellite loci mutate at an extremely high rate and are generally thought to evolve through a stepwise mutation model. Several differentiation statistics taking into account the particular mutation scheme of the microsatellite have been proposed. The most commonly used is $\widehat{R_{ST}}$, which is independent of the mutation rate under a generalized stepwise mutation model. $\widehat{F_{ST}}$ and $\widehat{R_{ST}}$ are commonly reported in the literature, but often differ widely. Here we compare their statistical performances using individual-based simulations of a finite island model. The simulations were run under different levels of gene flow, mutation rates, population number and sizes. In addition to the per locus statistical properties, we compare two ways of combining $\widehat{R_{ST}}$ over loci. Our simulations show that even under a strict stepwise mutation model, no statistic is best overall. All estimators suffer to different extents from large bias and variance. While $\widehat{R_{ST}}$ better reflects population differentiation in populations characterized by very low gene-exchange, $\widehat{F_{ST}}$ gives better estimates in cases of high levels of gene flow. The number of loci sampled (12, 24, or 96) has only a minor effect on the relative performance of the estimators under study. For all estimators there is a striking effect of the number of samples, with the differentiation estimates showing very odd distributions for two samples.

Keywords: computer simulations, F_{ST}, gene flow, microsatellites, population structure, R_{ST}

Received 14 October 2001; revision received 3 January 2002; accepted 3 January 2002

Introduction

Microsatellite markers have had a tremendous impact on population genetics. They have became the most commonly used co-dominant genetic markers. Microsatellites mutate at an extremely high rate and are generally believed to evolve mainly under a stepwise mutation scheme, characterized by the addition or deletion of one ore more repeated motifs (e.g. Weber & Wong 1993; Brinkmann *et al*. 1998; Di Rienzo *et al*. 1998; Xu *et al*. 2000).

The high mutation rate of microsatellites has drawbacks. The probability of identity (by descent or by state) of two genes decreases as the mutation rate increases (Rousset 1996). F_{ST}, a function of probabilities of identity, has a lowered expectation when the mutation rate is high (Wright 1978; Charlesworth 1998; Nagylaki 1998; Hedrick 1999; Balloux *et al*. 2000) and inferences drawn from F_{ST}, such as the number of migrants M, will be biased. This bias stems from the impossibility to separate the effects of mutation and migration. Fortunately, it has been shown that when the mutation process is stepwise (it could actually be multistep and nonsymmetrical, see Kimmel *et al*. 1996), mutation can be disentangled from processes such as migration (Slatkin 1995; Rousset 1996). However, statistics based on the probability of identity of alleles are not sufficient for this, and we need to account for the evolutionary distance between alleles. For markers evolving under a stepwise mutation model, size differences are related to evolutionary distances between alleles (Goldstein *et al*. 1995; Michalakis & Excoffier 1996). Several statistics accounting for the size of alleles have been developed for the estimation of genetic distances and population differentiation under stepwise mutation (e.g. Goldstein *et al*. 1995; Slatkin 1995). The statistic most commonly used for population differentiation is R_{ST} (originally defined by Chakraborty & Nei

Correspondence: Jérôme Goudet. Fax: + 41 21 692 41 05; E-mail: jerome.goudet@ie-zea.unil.ch

1982; independently derived by Slatkin 1995), which is independent of the mutation rate under a generalized stepwise mutation model. The drawback of R_{ST} is its high associated variance compared to other descriptors of population structure such as F_{ST} (Slatkin 1995; Balloux *et al.* 2000). Since both R_{ST} and F_{ST} are very commonly reported for studies using microsatellite markers, and often differ widely (reviewed in Lugon-Moulin *et al.* 1999), it seems of interest to compare their respective behaviour under varying sampling schemes.

Surprisingly there has been little work on the statistical properties of these differentiation estimators (other statistics have received more attention, see Kimmel & Chakraborty 1996; Pritchard & Feldman 1996). Slatkin (1995) reported some simulations based on two samples from 10 populations showing that under stepwise mutations, estimates $\hat{M}$ of the number of migrants (Nm) from $\widehat{R_{ST}}$ were less biased than those derived from $\widehat{F_{ST}}$. More recently Gaggiotti *et al.* (1999) compared the performance of $\hat{M}$ based on $\widehat{F_{ST}}$ [$\mathrm{M}(\widehat{F_{ST}})$] and $\widehat{R_{ST}}$[$\mathrm{M}(\widehat{R_{ST}})$] also in the case of two populations. They concluded that the relative performance of $\mathrm{M}(\widehat{F_{ST}})$ and $\mathrm{M}(\widehat{R_{ST}})$ was dependent mainly on sample size, with $\mathrm{M}(\widehat{F_{ST}})$ being a better estimator for small samples. But we are aware of no studies addressing specifically the statistical properties of estimators of F_{ST} and R_{ST}.

Here, using computer simulations, we present an extensive comparison of the statistical properties of estimators of the parameters F_{ST} and R_{ST} [actually the θ of Weir & Cockerham (1984) and the ρ_{ST} of Rousset (1996)] in a finite island model under different levels of gene flow, mutation rates, population number and sizes. The simulations are restricted to a strict symmetrical stepwise mutation model. In addition to exploring properties of statistics for individual loci, we compare two methods for combining estimates across loci. *F*- and *R*-statistics are ratios of variances. There is no consensus in the statistical literature concerning the best way to estimate these ratios, as one could take the ratio of the averages or the average over the ratios (see for instance King *et al.* 2000). For $\widehat{F_{ST}}$, simulations in Weir & Cockerham (1984) showed that the best way to combine information across loci depends on the level of differentiation. However, most studies report the ratio of averages, as was suggested by Weir & Cockerham (1984) and Weir (1996). For *R*-statistics, two solutions are advocated in the literature. Goldstein *et al.* (1995), Slatkin (1995) and Michalakis & Excoffier (1996) use the ratio of the averages while Goodman (1997) suggested standardizing the variance of allele size prior to calculation. In this way all loci are given the same weight independently of their variance. This seems to be intuitively reasonable since microsatellite loci often show manifold differences in allele-size variance. Last, we investigate the consequences of subsampling the simulations, focalizing on samples of 20 individuals from two demes.

Materials and methods

Simulations

We used individual-based simulations to assess the statistical properties of each estimator. The software EASYPOP (version 1.7) (Balloux 2001) was used to generate populations of different structures. In all cases, we simulated populations made of a fixed number of 2000 individuals. These individuals were either arranged in two demes of 1000 individuals, five demes of 400 individuals or 20 demes of 100 individuals. Migration of individuals (rather than gametes) among demes followed the island model of migration and the number of migrants was fixed to 0.1, 1, or 10. For each replicate, 12 loci were simulated. Replicates were run for mutation rates μ of 10^{-2}, 10^{-3} and 10^{-4} as well as a mixed situation where four loci were set at 10^{-2}, four at 10^{-3} and four at 10^{-4}. Mutations followed a single-step mutation model with 999 possible allelic states. With this large number of alleles, the mutation model can be considered as unconstrained (Balloux *et al.* 2000). The simulations were run for 10 000 generations ($5nN$ generations, where n is the number of demes and N is the size of each deme), point at which all statistics had reached equilibrium, and replicated 99 times. Differentiation statistics were computed from the final generation of each simulation using the software FSTAT version 2.9.3 (Goudet 2001, updated from Goudet 1995). Single locus $\widehat{F_{ST}}$ and $\widehat{R_{ST}}$ were estimated following Weir & Cockerham (1984) and Rousset (1996), respectively.

Expectations of F_{ST} and R_{ST} under stepwise mutation model

In order to assess the relative performance of the differentiation estimators under study, and to quantify their bias, we need their parametric values. Since all simulations have been performed under a stepwise mutation model, the theoretical value of R_{ST} (E[R_{ST}]) is a function of gene flow only (Slatkin 1995; Rousset 1996). We will therefore estimate the bias of $\widehat{F_{ST}}$ and $\widehat{R_{ST}}$ using E[R_{ST}] as the reference. Mean square errors (MSE; Rice 1995), computed as the sum of the squared bias and the variance (Bias^2 + Var), were used to quantify the efficiency of the respective statistics. Theoretical values for R_{ST} and F_{ST}, variance in allele sizes and gene diversity for our model of population structure are given in Appendix I.

Combining estimates across loci

Multilocus $\widehat{F_{ST}}$ and $\widehat{R_{ST}}$ were computed as the ratio of the sum (over loci) of variances following Weir & Cockerham (1984) and Michalakis & Excoffier (1996):

$$\widehat{F_{ST}}, \widehat{R_{ST\mathrm{w}}} = \frac{\sum \hat{X}_a}{\sum \hat{X}_t} \qquad (1)$$

where $\hat{X}$ stands for either the variance components of allele frequencies ($\widehat{F_{ST}}$) or the variances of allele size ($\widehat{R_{ST}}$), the index a stands for among populations and t represents the overall population. Alternatively, $\widehat{R_{ST}}$ can be computed as the average of ratios:

$$\widehat{R_{STu}} = \frac{1}{l}\sum\frac{\hat{V}_a}{\hat{V}_t} \qquad (2)$$

where l stands for the number of scored loci. We show in Appendix II that $\widehat{R_{STu}}$ is similar to the value $\widehat{UR_{ST}}$ in Goodman (1997).

To investigate how increasing the number of loci would affect the variance of the different estimators, we pooled replicates to obtain 49 simulations with 24 loci and 12 simulations with 96 loci. We checked that there was no difference in the components of variances among replicates using a one-way analysis of variance (ANOVA) after rank transformation of the variance components.

Sub-sampling simulations

Empirical population genetics datasets rarely comprise the exhaustive population under study. To explore the effect of nonexhaustive sampling, we subsampled from the simulated populations. Our goal is not to investigate extensively the effect of partial sampling strategies, but to explore how our conclusions would be affected by sampling only a small part of the population. We sampled randomly 20 individuals in two populations, a sampling scheme similar to that used by Slatkin (1995) and Gaggiotti *et al*. (1999).

Results

Theoretical values of F$_{ST}$ *and* R$_{ST}$

Table 1 shows the expected values of F_{ST} and R_{ST} for the different simulation scenarios. While for a given Nm, E[R_{ST}] changes with the number of populations, it remains constant for varying mutation rates. On the other hand, E[F_{ST}] varies wildly with mutation rates, particularly for low migrant numbers. E[F_{ST}] is always lower than E[R_{ST}], and in the extreme case of low Nm and high mutation rates, E[R_{ST}] is over eight times as large. E[R_{ST}] and E[F_{ST}] tend to similar values as migration increases (Table 1).

Single locus estimates

The results for individual loci statistics are given in Table 1 and Fig. 1. Average $\widehat{F_{ST}}$ over all simulations lie close to their expectation most of the time, but very far from E[R_{ST}] when migration is low (Fig. 1e–h and Table 1). It is only under very low mutation rate and with only two populations that average $\widehat{F_{ST}}$ strongly underestimate E[F_{ST}] (20% bias, Table 1). Average $\widehat{R_{ST}}$ values also lie far from their own expectation when migration is low and number of populations are small (Fig. 1a,c and Table 1). $\widehat{R_{ST}}$ average values improve with increasing migration (Table 1). The number of populations has a striking effect since the distributions of both individual loci $\widehat{F_{ST}}$ and $\widehat{R_{ST}}$ display a much larger variance for two than for 20 populations (Fig. 1). Average $\widehat{R_{ST}}$ values are downwardly biased, but this effect tends to disappear as the number of populations increases (Table 1). The distributions of $\widehat{R_{ST}}$ are surprising. With two populations, they are completely asymmetrical with a skew on the left and a mode around 0 (Fig. 1a,c). There is no effect of the mutation rate on $\widehat{R_{ST}}$. This was expected for the mean, but it appears to be true for all higher moments of the distribution. There is a strong effect of mutation rates on the mean and distribution of $\widehat{F_{ST}}$, and it is particularly affected by a heterogeneous mutation rate (Fig. 1e–h). This was expected, as the effect of migration cannot be disentangled from that of mutation for statistics based on the probability of identity. We also note that the distribution of individual loci $\widehat{F_{ST}}$ for very low gene flow and mutation rate looks very similar to the distribution of individual loci $\widehat{R_{ST}}$ with very low gene flow (data not shown).

Averaging over loci

Behaviour of the statistics under exhaustive sampling. Because $\widehat{F_{ST}}$ estimated as the ratio of sums consistently showed lower MSE than its counterpart estimated as the average of ratios (data not shown), we will only report multilocus $\widehat{F_{ST}}$ as defined in equation 1.

Figure 2 shows the distribution of $\widehat{R_{STu}}$, $\widehat{R_{STw}}$ and $\widehat{F_{ST}}$ for 0.1 migrants per generation, where the statistics are estimated over 12 loci. With this number of loci, the distribution of the different statistics is slightly more bell shaped than when each locus is taken individually. The distribution of $\widehat{R_{STw}}$ retains a very large variance (Fig. 2, second row) but its average value is dramatically improved (Table 1). When the simulation consists of two populations (left hand side of Fig. 2), $\widehat{R_{STu}}$ is biased strongly downward (Fig. 2a and Table 1). On the other hand, $\widehat{R_{STu}}$ has a smaller variance than $\widehat{R_{STw}}$. As the number of populations increases (right-hand side of Fig. 2 and Table 1), bias and variance of the two $\widehat{R_{ST}}$ are reduced (note the change of scale on the x-axis). The distribution of $\widehat{F_{ST}}$ (last row of Fig. 2) is well centred around its expected value (dashed line, see also Table 1), but very much offset to the left of the expected value of R_{ST} (solid line). The variance of $\widehat{F_{ST}}$ also diminishes as the number of populations increases.

Figures 3 and 4 display the distribution of the same statistics for a number of migrants per generation of 1 and 10, respectively. As migration increases, bias and variance diminish. $\widehat{R_{STu}}$ remains more biased but with a smaller

Table 1 Average values of differentiation statistics as a function of the number of scored loci

					Expectations		1-locus estimate		12-locus estimate			96-locus estimate		
n	N	m	μ	Nm	E[R_{ST}]	E(F_{ST})	$\widehat{R_{ST}}$	$\widehat{F_{ST}}$	$\widehat{R_{STu}}$	$\widehat{R_{STw}}$	$\widehat{F_{ST}}$	$\widehat{R_{STu}}$	$\widehat{R_{STw}}$	$\widehat{F_{ST}}$
2	1000	0.0001	0.01	0.1	0.556	0.065	0.374	0.065	0.374	0.522	0.065	0.376	0.564	0.066
5	400	0.00025	0.01	0.1	0.667	0.122	0.604	0.121	0.604	0.661	0.121	0.605	0.667	0.121
20	100	0.001	0.01	0.1	0.704	0.255	0.694	0.255	0.694	0.702	0.255	0.694	0.701	0.255
2	1000	0.0001	0.001	0.1	0.556	0.202	0.390	0.201	0.390	0.520	0.208	0.390	0.543	0.208
5	400	0.00025	0.001	0.1	0.667	0.344	0.603	0.342	0.603	0.654	0.344	0.603	0.661	0.344
20	100	0.001	0.001	0.1	0.704	0.547	0.692	0.545	0.692	0.702	0.546	0.692	0.703	0.546
2	1000	0.0001	0.0001	0.1	0.556	0.434	0.356	0.338	0.359	0.520	0.434	0.361	0.546	0.442
5	400	0.00025	0.0001	0.1	0.667	0.588	0.597	0.561	0.599	0.672	0.591	0.600	0.672	0.593
20	100	0.001	0.0001	0.1	0.704	0.680	0.693	0.678	0.693	0.710	0.683	0.693	0.710	0.683
2	1000	0.0001	mix	0.1	0.556	0.202	0.358	0.192	0.358	0.490	0.194	0.358	0.534	0.195
5	400	0.00025	mix	0.1	0.667	0.316	0.598	0.340	0.598	0.639	0.314	0.597	0.642	0.315
20	100	0.001	mix	0.1	0.704	0.465	0.693	0.495	0.693	0.702	0.467	0.693	0.706	0.467
2	1000	0.001	0.01	1	0.111	0.031	0.097	0.031	0.097	0.119	0.031	0.097	0.124	0.031
5	400	0.0025	0.01	1	0.166	0.060	0.158	0.060	0.158	0.166	0.060	0.157	0.168	0.060
20	100	0.01	0.01	1	0.190	0.114	0.188	0.114	0.188	0.192	0.114	0.188	0.191	0.114
2	1000	0.001	0.001	1	0.111	0.075	0.092	0.072	0.092	0.101	0.074	0.092	0.102	0.074
5	400	0.0025	0.001	1	0.166	0.128	0.154	0.126	0.154	0.161	0.127	0.154	0.161	0.127
20	100	0.01	0.001	1	0.190	0.175	0.187	0.174	0.187	0.188	0.174	0.187	0.189	0.174
2	1000	0.001	0.0001	1	0.111	0.104	0.078	0.081	0.080	0.094	0.096	0.079	0.096	0.098
5	400	0.0025	0.0001	1	0.166	0.160	0.144	0.145	0.144	0.166	0.160	0.144	0.166	0.160
20	100	0.01	0.0001	1	0.190	0.188	0.181	0.181	0.181	0.191	0.188	0.181	0.192	0.188
2	1000	0.001	mix	1	0.111	0.061	0.090	0.065	0.090	0.116	0.063	0.089	0.124	0.063
5	400	0.0025	mix	1	0.166	0.105	0.149	0.110	0.149	0.150	0.105	0.149	0.147	0.105
20	100	0.01	mix	1	0.190	0.151	0.182	0.155	0.183	0.186	0.149	0.182	0.187	0.149
2	1000	0.01	0.01	10	0.012	0.008	0.012	0.008	0.012	0.011	0.008	0.012	0.011	0.008
5	400	0.025	0.01	10	0.019	0.014	0.019	0.014	0.019	0.019	0.014	0.019	0.019	0.014
20	100	0.1	0.01	10	0.020	0.018	0.020	0.018	0.020	0.019	0.018	0.020	0.019	0.018
2	1000	0.01	0.001	10	0.012	0.011	0.012	0.011	0.012	0.010	0.011	0.012	0.010	0.011
5	400	0.025	0.001	10	0.019	0.018	0.019	0.018	0.019	0.018	0.018	0.019	0.018	0.018
20	100	0.1	0.001	10	0.020	0.020	0.020	0.020	0.020	0.018	0.020	0.020	0.018	0.020
2	1000	0.01	0.0001	10	0.012	0.012	0.011	0.011	0.011	0.012	0.012	0.012	0.013	0.012
5	400	0.025	0.0001	10	0.019	0.019	0.018	0.018	0.018	0.019	0.019	0.018	0.019	0.019
20	100	0.1	0.0001	10	0.020	0.020	0.019	0.019	0.019	0.020	0.020	0.019	0.020	0.020
2	1000	0.01	mix	10	0.012	0.010	0.012	0.010	0.012	0.013	0.010	0.012	0.013	0.010
5	400	0.025	mix	10	0.019	0.016	0.018	0.016	0.018	0.018	0.016	0.018	0.018	0.016
20	100	0.1	mix	10	0.020	0.019	0.020	0.019	0.020	0.020	0.019	0.020	0.020	0.019

n, the number of populations; N, the number of individuals per population; m, the migration rate; μ, the mutation rate; E[R_{ST}] stands for the expectation of R_{ST} and E[F_{ST}] stands for the expectation of F_{ST}. $\widehat{R_{STw}}$ is the multilocus estimator of R_{ST} according to equation 1 in the text, while $\widehat{R_{STu}}$ is the multilocus estimator of R_{ST} according to equation 2. The average for single locus estimates is based on 1188 replicates, that for 12 loci estimates is based on 99 replicates, while the average for 96 loci estimates is based on 12 replicates only.

variance than $\widehat{R_{STw}}$, and these two characteristics of the distribution decrease with the number of populations. $\widehat{F_{ST}}$ has a much smaller variance than the two $\widehat{R_{ST}}$ statistics, and its distribution is well centred around its expectation (last row of Figs 3 and 4). But $\widehat{F_{ST}}$ strongly underestimates differentiation when compared to the expectation of R_{ST}. Figure 4 shows that when dispersal is large (10 migrants per generation), the variance of $\widehat{R_{STw}}$ remains larger than that of the two other statistics. It is notable that Fig. 4(b) shows the small variance and bias of $\widehat{R_{STu}}$ when the number of population is large. $\widehat{F_{ST}}$ also has a small variance in this case, but gives a slight underestimation of the expectation of R_{ST} (Fig. 4f).

The joint effect of bias and variance is summarized by the MSE, shown in Table 2 as its square root for convenience. Whatever the statistic, MSE decreases as the number of migrants increases. MSE also decreases dramatically as the number of populations increases, in particular for $\widehat{R_{STw}}$ and $\widehat{R_{STu}}$ and low gene flow. This information is summarized in the column 'best statistic' of Table 2, where best is

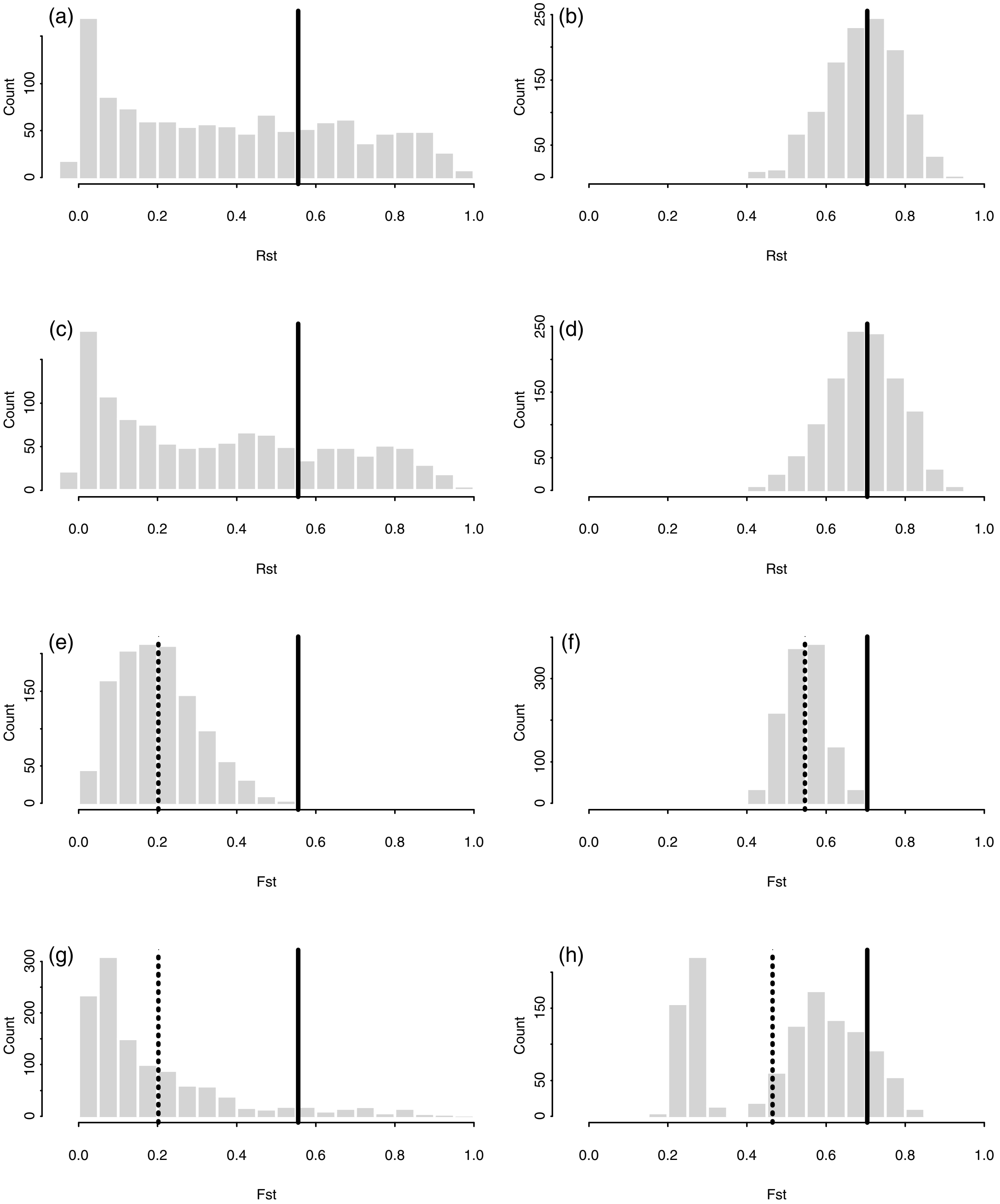

Fig. 1 Single locus distribution of $\widehat{R_{ST}}$ and $\widehat{F_{ST}}$ for a number of migrants $Nm = 0.1$. On all panels, the solid black vertical line represents the expectation of R_{ST}, while the dotted line represent the expectation of F_{ST}. Panels (a) to (d): distribution of $\widehat{R_{ST}}$. Panels (e) to (h): distribution of $\widehat{F_{ST}}$. (a) and (e) two populations, $\mu = 10^{-3}$; (b) and (f) 20 populations, $\mu = 10^{-3}$; (c) and (g) two populations, mixed mutation rate; (d) and (h) 20 populations, mixed mutation rate.

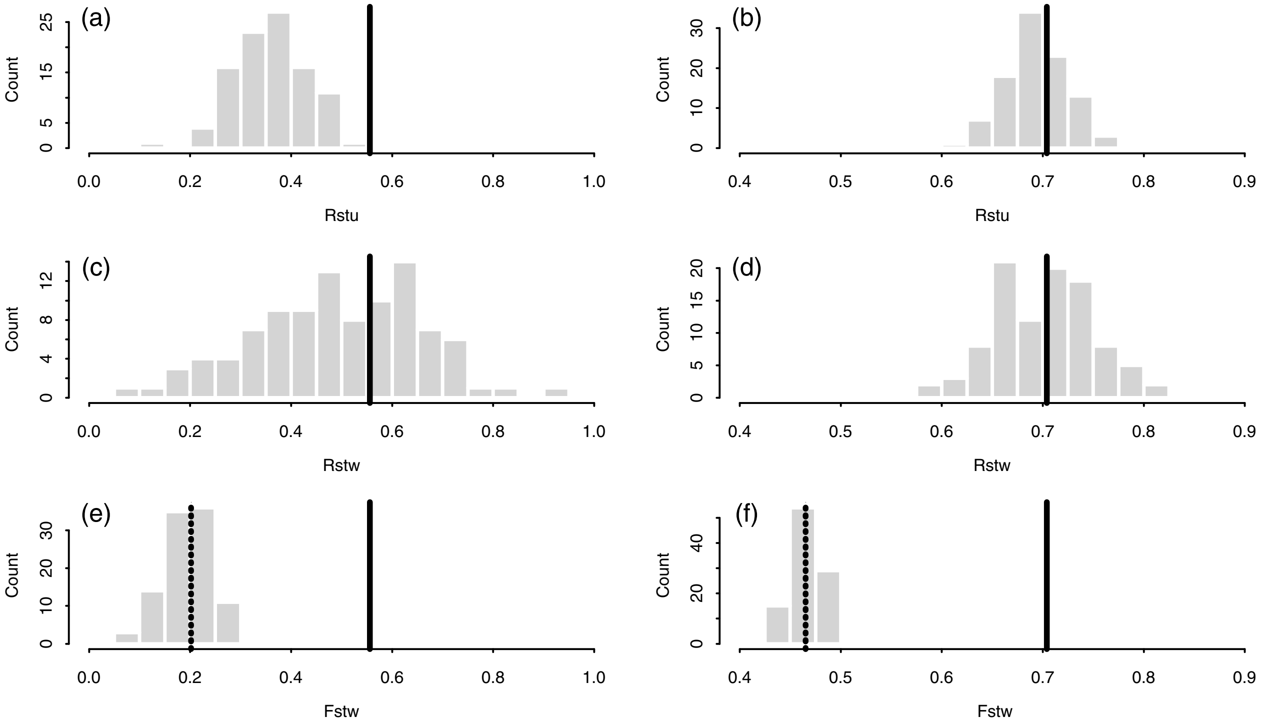

Fig. 2 Distribution of $\widehat{R_{STu}}$, $\widehat{R_{STw}}$ and $\widehat{F_{ST}}$ for a number of migrants, *Nm* equal to 0.1. Each estimate is based on 12 loci. The mutation rate is mixed (four loci at 10^{-4}, four at 10^{-3} and four at 10^{-2}). Panels (a) and (b): distribution of $\widehat{R_{STu}}$; panels (c) and (d): distribution of $\widehat{R_{STw}}$; panels (e) and (f): distribution of $\widehat{F_{ST}}$. Panels (a) (c) and (e): two populations; panels (b) (d) and (f): 20 populations. The vertical solid line represents the expectation of R_{ST}, the vertical dotted line that of F_{ST}.

Table 2 Mean Square Error (MSE) of the three statistics, $\widehat{F_{ST}}$, $\widehat{R_{STu}}$ and $\widehat{R_{STw}}$ for estimates based on 12, 24 and 96 loci

				Expectations		Sq root MSE $\widehat{F_{ST}}$ (×100)			Sq root MSE $\widehat{R_{STu}}$ (×100)			Sq root MSE $\widehat{R_{STw}}$ (×100)			Best statistic		
n	*N*	*m*	μ	E(R_{ST})	E(F_{ST})	12 loci	24 loci	96 loci	12 loci	24 loci	96 loci	12 loci	24 loci	96 loci	12 loci	24 loci	96 loci
2	1000	0.0001	0.001	0.556	0.202	34.9	34.9	34.8	18.1	17.4	16.7	12.6	9.3	5.2	*Rw*	*Rw*	*Rw*
5	400	0.00025	0.001	0.667	0.344	32.4	32.3	32.3	8.4	7.4	6.5	6.7	5.2	2.2	*Rw*	*Rw*	*Rw*
20	100	0.001	0.001	0.704	0.547	15.9	15.9	15.9	2.9	2.1	1.4	3.6	2.5	1.3	*Ru*	*Ru*	*Rw*
2	1000	0.0001	mix	0.556	0.202	36.5	36.3	36.0	21.0	20.3	20.0	17.7	13.5	7.7	*Rw*	*Rw*	*Rw*
5	400	0.00025	mix	0.667	0.316	35.4	35.3	35.2	8.8	7.7	7.2	10.2	6.8	3.9	*Ru*	*Rw*	*Rw*
20	100	0.001	mix	0.704	0.465	23.8	23.8	23.7	3.3	2.3	1.6	4.9	3.0	1.9	*Ru*	*Ru*	*Ru*
2	1000	0.001	0.001	0.111	0.075	4.0	3.8	3.7	3.5	2.7	2.0	4.5	3.3	2.1	*Ru*	*Ru*	*Ru*
5	400	0.0025	0.001	0.166	0.128	4.1	4.0	4.0	2.9	2.1	1.7	3.7	2.8	1.7	*Ru*	*Ru*	*Rw*
20	100	0.01	0.001	0.190	0.175	1.9	1.8	1.6	1.7	1.2	0.6	1.9	1.5	0.8	*Ru*	*Ru*	*Ru*
2	1000	0.001	mix	0.111	0.061	5.0	4.9	4.9	3.7	3.1	2.4	7.9	6.5	2.9	*Ru*	*Ru*	*Ru*
5	400	0.0025	mix	0.166	0.105	6.2	6.2	6.1	3.2	2.5	2.1	4.8	3.3	2.3	*Ru*	*Ru*	*Ru*
20	100	0.01	mix	0.190	0.151	4.2	4.1	4.1	2.1	1.5	1.2	2.9	2.2	1.1	*Ru*	*Ru*	*Rw*
2	1000	0.01	0.001	0.012	0.011	0.2	0.2	0.1	0.4	0.3	0.1	0.7	0.6	0.3	*Fw*	*Fw*	*Ru*
5	400	0.025	0.001	0.019	0.018	0.2	0.2	0.1	0.5	0.3	0.1	0.7	0.5	0.3	*Fw*	*Fw*	*Fw*
20	100	0.1	0.001	0.020	0.020	0.1	0.1	0.0	0.3	0.2	0.1	0.4	0.3	0.2	*Fw*	*Fw*	*Fw*
2	1000	0.01	mix	0.012	0.010	0.3	0.3	0.2	0.5	0.3	0.2	0.9	0.6	0.4	*Fw*	*Fw*	*Ru*
5	400	0.025	mix	0.019	0.016	0.3	0.3	0.3	0.4	0.3	0.2	0.8	0.6	0.4	*Fw*	*Fw*	*Ru*
20	100	0.1	mix	0.020	0.019	0.2	0.1	0.1	0.2	0.1	0.1	0.5	0.3	0.2	*Fw*	*Ru*	*Ru*

For convenience and clarity of reading, the square root of MSE (×100) is represented rather than MSE. The last three columns, labelled 'Best Statistic' show which of the three statistics has the lowest MSE. *Rw* stands for $\widehat{R_{STw}}$, *Ru* stands for $\widehat{R_{STu}}$ and *Fw* stands for $\widehat{F_{ST}}$.

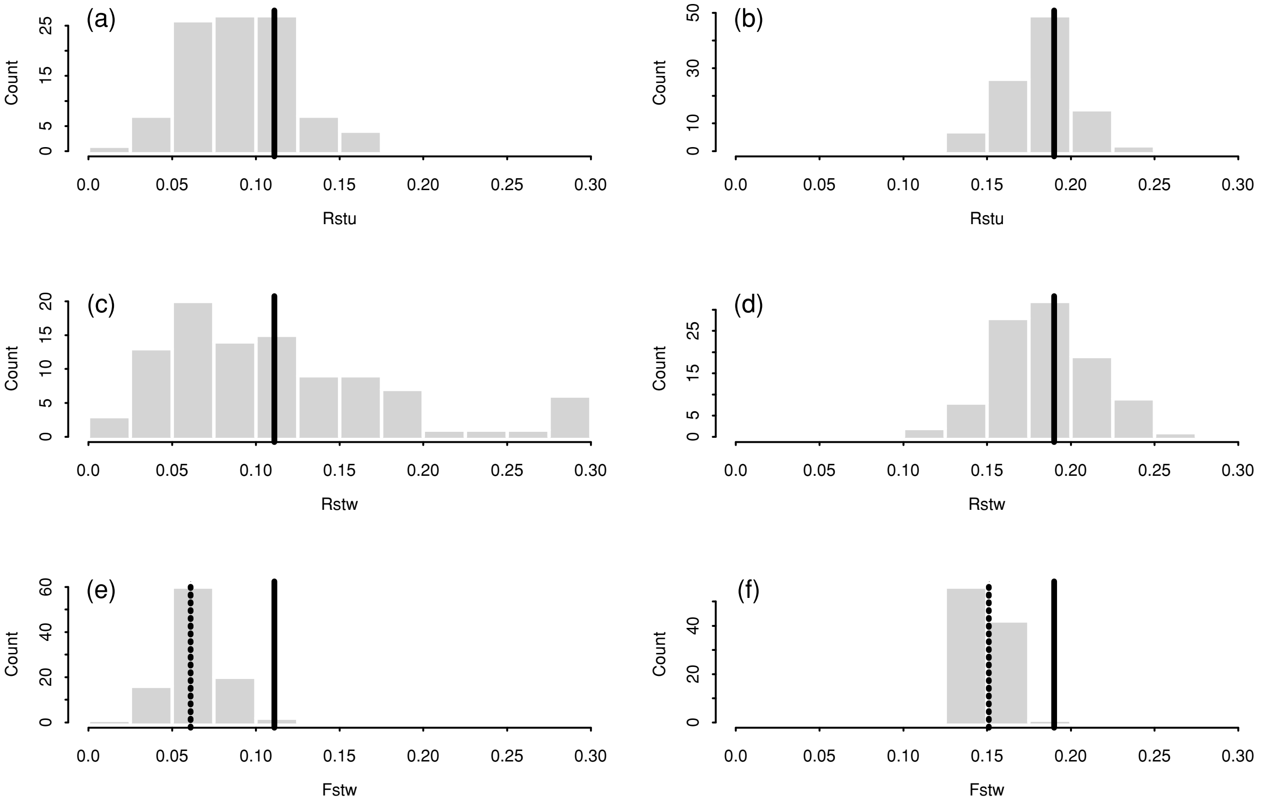

Fig. 3 Distribution of $\widehat{R_{STu}}$, $\widehat{R_{STw}}$ and $\widehat{F_{ST}}$ for a number of migrants, *Nm* equal to 1. Each estimate is based on 12 loci. The mutation rate is mixed (four loci at 10^{-4}, four at 10^{-3} and four at 10^{-2}). Panels (a) and (b): distribution of $\widehat{R_{STu}}$; panels (c) and (d): distribution of $\widehat{R_{STw}}$; panels (e) and (f): distribution of $\widehat{F_{ST}}$. Panels (a) (c) and (e): two populations; panels (b) (d) and (f): 20 populations. The vertical solid line represents the expectation of R_{ST}, the vertical dotted line that of F_{ST}.

used with the statistical sense of efficiency (see Rice 1995). $\widehat{R_{STw}}$ is the best statistic to use for low migration rate and small number of populations. As gene flow or the number of populations increases, $\widehat{R_{STu}}$ becomes the statistic of choice. For the highest rate of gene flow (10 migrants per generation) the best statistic is $\widehat{F_{ST}}$. For smaller mutation rates, $\widehat{F_{ST}}$ also tends to outperform $\widehat{R_{STu}}$ for moderate migration rates (data not shown).

Effect of the number of loci. Table 1 shows that the average $\widehat{R_{STw}}$ improves as we move from 12 to 96 loci, particularly for low *Nm* and small number of populations. However, this improvement is slight. The changes for $\widehat{F_{ST}}$ are minute. Table 2 shows the effect of increasing the number of loci to 24 and 96 on MSE. This increase in the number of loci should affect the variance of the statistic, but not its bias. Therefore, we expect statistics with high variance and low bias to become better. Indeed, the statistic with the smallest bias but the largest variance, $\widehat{R_{STw}}$, benefits most from an increased number of loci, followed by $\widehat{R_{STu}}$ (Table 2). $\widehat{F_{ST}}$ benefits close to nothing from this increased number of loci.

However, from the columns labelled 'best statistic' in Table 2 it is clear that doubling or even multiplying by eight the number of typed loci has very little effect on which statistic is best. When there is a change, it follows expectations, the statistic with small bias but large variance (e.g. $\widehat{R_{STw}}$ compared to $\widehat{R_{STu}}$, or $\widehat{R_{STu}}$ compared to $\widehat{F_{ST}}$) becoming better. Overall, there is not a best statistic for all situations. While mutation rate heterogeneity increases MSE slightly, it does not affect which statistic is best.

Partial sampling. The results for partial sampling are given in Table 3. The statistics affected most by the sampling scheme are those with large variances. $\widehat{R_{STw}}$ deteriorates most, followed by $\widehat{R_{STu}}$, while $\widehat{F_{ST}}$ is little affected.

When the original simulation consists of two populations, MSE changes little (compare Tables 2 and 3, first row for instance). The largest increases in MSE appear when subsamples are taken from a population originally comprising 20 populations (e.g. the third row of Tables 2 and 3). This increase goes from fourfold for low migration to 10-fold for the highest migration. Although when sampling exhaustively MSE decreases as the number of populations increases (Table 2), the reverse is true under a partial sampling scheme of 2 × 20 (Table 3).

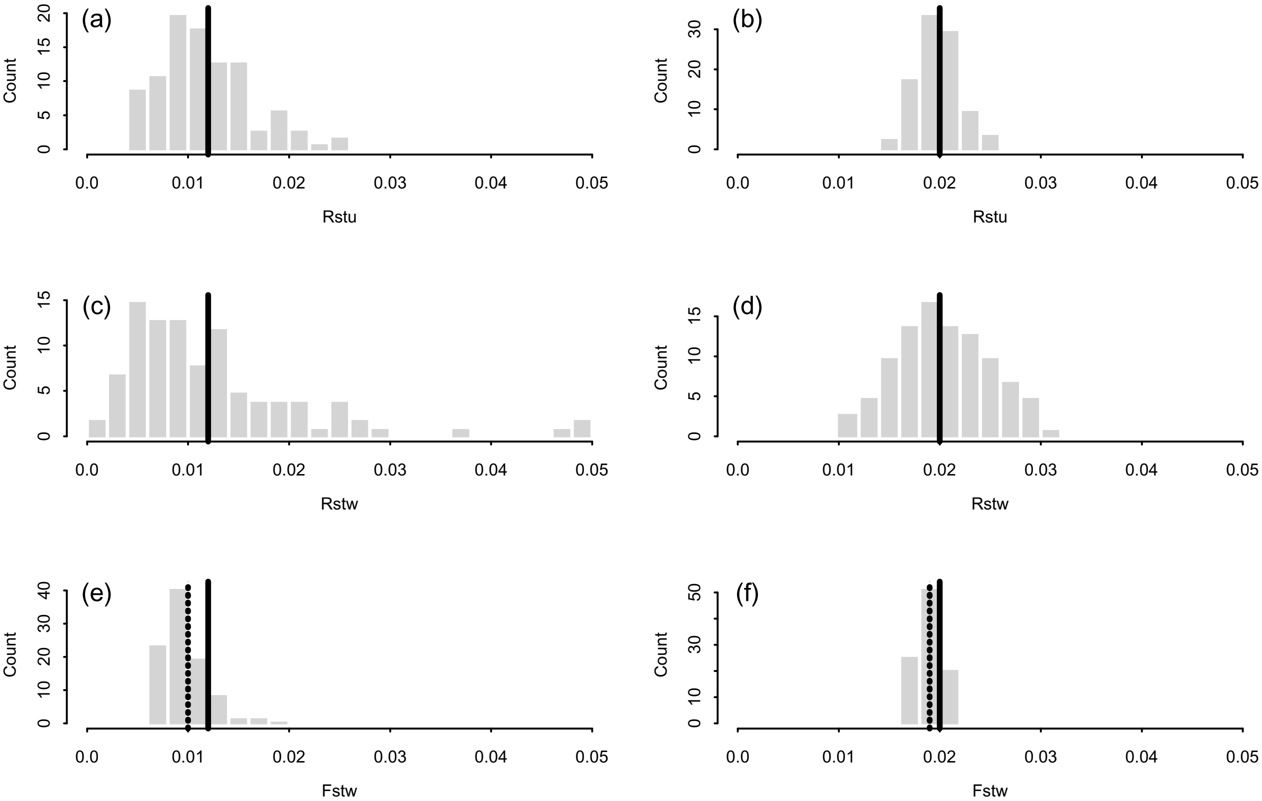

Fig. 4 Distribution of $\widehat{R_{\mathrm{STu}}}$, $\widehat{R_{\mathrm{STw}}}$ and $\widehat{F_{\mathrm{ST}}}$ for a number of migrants, *Nm* equal to 10. Each estimate is based on 12 loci. The mutation rate is mixed (four loci at 10^{-4}, four at 10^{-3} and four at 10^{-2}). Panels (a) and (b): distribution of $\widehat{R_{\mathrm{STu}}}$; panels (c) and (d): distribution of $\widehat{R_{\mathrm{STw}}}$; panels (e) and (f): distribution of $\widehat{F_{\mathrm{ST}}}$. Panels (a) (c) and (e): two populations; panels (b) (d) and (f): 20 populations. The vertical solid line represents the expectation of R_{ST}, the vertical dotted line that of F_{ST}.

Discussion

We have shown that under restricted gene flow and a pure stepwise mutation model, there is not a single best estimator of population subdivision. The existing estimators suffer from large bias and large variance to different extents. When populations are highly structured ($Nm = 0.1$), $\widehat{R_{\mathrm{STw}}}$ does best, particularly when the number of samples is small. When populations are very weakly structured ($Nm = 10$), $\widehat{F_{\mathrm{ST}}}$ does best. For intermediate situation ($Nm = 1$), $\widehat{R_{\mathrm{STu}}}$ is the statistic of choice, although when very small samples are taken, $\widehat{F_{\mathrm{ST}}}$ does slightly better. Although increasing the number of loci decreases MSE for all statistics under study, our simulations show that the 'best' statistic is almost independent of the number of typed loci (Tables 2 and 3). These conclusions are valid for the intermediate to large mutation rates commonly found in empirical surveys of microsatellites (corresponding to gene diversities larger than 70%). For lower mutation rates (leading to gene diversities of 50% or lower), the tendency is for $\widehat{F_{\mathrm{ST}}}$ to do better than $\widehat{R_{\mathrm{ST}}}$ (data not shown). However, as mutation rate decreases, the distribution of $\widehat{F_{\mathrm{ST}}}$ (and therefore its mean and variance) becomes similar to that of $\widehat{R_{\mathrm{ST}}}$.

Generally, these results are concordant with the empirical findings reviewed by Lugon-Moulin *et al.* (1999). These authors reviewed empirical studies based on microsatellites, where $\widehat{F_{\mathrm{ST}}}$ and $\widehat{R_{\mathrm{ST}}}$ had been obtained. For moderate to strong differentiation, $\widehat{R_{\mathrm{ST}}}$ seems a better estimator, because it was larger than $\widehat{F_{\mathrm{ST}}}$ in 13 studies out of the 15 reviewed. While there is some circularity in this argument, since $\widehat{R_{\mathrm{ST}}}$ is used to define the level of structuring, the geographical level at which these studies were carried out is compatible with little genetic exchange. For low levels of differentiation, these authors found that estimates of $\widehat{F_{\mathrm{ST}}}$ were larger than $\widehat{R_{\mathrm{ST}}}$ in six studies out of eight. This pattern is not expected (as differentiation decreases, the expectations of $\widehat{F_{\mathrm{ST}}}$ and $\widehat{R_{\mathrm{ST}}}$ converge, but $\widehat{F_{\mathrm{ST}}}$ never becomes larger than $\widehat{R_{\mathrm{ST}}}$, even under mutation models others than a stepwise mutation model, unless allele size differences are inversely related to evolutionary distances between alleles), but might be explained by the higher variance of $\widehat{R_{\mathrm{ST}}}$ estimators. On the whole therefore, empirical estimates are in agreement with our simulation results concerning the ranking of the different statistics.

Before discussing the empirical evidence for a stepwise mutation model, we first address the issue of estimating gene flow from estimators of population differentiation. As a start we note that methods other than those based on

Table 3 Mean Square Error (MSE) of the three statistics, $\widehat{F_{ST}}$, $\widehat{R_{STu}}$ and $\widehat{R_{STw}}$ for estimates based on 12 and 24 loci but for a subsample of each simulated dataset, consisting of two samples of 20 individuals

						Square root MSE (×100)						Best statistic	
n	N	m	μ	E[R_{ST}]	E[F_{ST}]	$\widehat{F_{ST}}$ 12 loci	$\widehat{F_{ST}}$ 24 loci	$\widehat{R_{STu}}$ 12 loci	$\widehat{R_{STu}}$ 24 loci	$\widehat{R_{STw}}$ 12 loci	$\widehat{R_{STw}}$ 24 loci	12 loci	24 loci
2	1000	0.0001	0.001	0.556	0.202	35.2	35.2	18.9	18.1	12.8	9.4	*Rw*	*Rw*
5	400	0.00025	0.001	0.667	0.344	32.7	32.6	19.6	18.5	13.9	11.6	*Rw*	*Rw*
20	100	0.001	0.001	0.704	0.547	18.1	17.5	21.7	20.6	16.0	10.9	*Rw*	*Rw*
2	1000	0.0001	mix	0.556	0.202	36.5	36.2	21.3	20.5	17.3	12.9	*Rw*	*Rw*
5	400	0.00025	mix	0.667	0.316	36.6	36.2	22.1	20.9	18.0	15.4	*Rw*	*Rw*
20	100	0.001	mix	0.704	0.465	24.9	24.3	20.0	17.8	20.8	15.5	*Ru*	*Rw*
2	1000	0.001	0.001	0.111	0.075	4.2	4.0	4.8	3.8	5.8	4.5	*Fw*	*Ru*
5	400	0.0025	0.001	0.166	0.128	5.1	4.6	6.3	5.2	7.8	6.2	*Fw*	*Fw*
20	100	0.01	0.001	0.190	0.175	4.5	3.4	6.1	5.0	8.3	6.3	*Fw*	*Fw*
2	1000	0.001	mix	0.111	0.061	5.0	4.9	4.7	3.7	8.9	6.9	*Ru*	*Ru*
5	400	0.0025	mix	0.166	0.105	7.0	6.7	6.0	5.0	10.1	7.4	*Ru*	*Ru*
20	100	0.01	mix	0.190	0.151	5.3	4.4	6.6	5.5	11.2	8.9	*Fw*	*Fw*
2	1000	0.01	0.001	0.012	0.011	0.8	0.6	1.5	1.1	1.6	1.3	*Fw*	*Fw*
5	400	0.025	0.001	0.019	0.018	0.9	0.6	1.6	1.1	2.5	1.7	*Fw*	*Fw*
20	100	0.1	0.001	0.020	0.020	1.0	0.8	1.5	1.0	2.2	1.5	*Fw*	*Fw*
2	1000	0.01	mix	0.012	0.010	0.7	0.5	1.4	1.0	2.6	2.1	*Fw*	*Fw*
5	400	0.025	mix	0.019	0.016	0.9	0.6	1.5	1.0	3.1	2.2	*Fw*	*Fw*
20	100	0.1	mix	0.020	0.019	0.9	0.7	1.7	1.3	3.7	3.0	*Fw*	*Fw*

F- and *R*-statistics have been recently developed to estimate gene flow, and often seem to perform better (see for instance Pritchard *et al*. 2000). Our purpose here is not to advocate the use of differentiation estimators to infer levels of gene flow (see Whitlock & McCauley 1999). Rather it is to clarify the recent reports in this journal of strongly biased estimates of the number of migrants M from $\widehat{R_{STw}}$ or $\widehat{F_{ST}}$ (Gaggiotti *et al*. 1999), and which might seem surprising in the light of our results. Indeed, Table 1 shows that the $\widehat{F_{ST}}$ of Weir & Cockerham (1984) is essentially unbiased. This is also true for $\widehat{R_{STw}}$ estimates based on 12 loci. Applying the classical relation:

$$M(i) \approx \frac{(1/i) - 1}{4n/(n-1)} \qquad (3)$$

(where i is one of $\widehat{F_{ST}}$ or $\widehat{R_{STw}}$) to results in Table 1 would therefore lead to unbiased estimates of M as long as migration was not too large. This is expected, since low migration is necessary for equation 3 to be valid (see Cockerham & Weir 1993). Even when subsampling 20 individuals in two demes, estimates of M from $\widehat{R_{STw}}$ always remain within 30% of their expectations (data not shown). The main reason for the difference between our simulations and those of Gaggiotti *et al*. (1999) stems from them averaging M estimates rather than differentiation estimates. As Cockerham & Weir (1993) pointed out (see equation 7 of their paper), M obtained from averaging M estimates will be inflated compared to a situation where differentiation estimates are first averaged, and then equation 3 is applied. Intuitively, this is because small variations on $\widehat{F_{ST}}$ (or $\widehat{R_{ST}}$) when differentiation is small translate into very large differences for M estimates, because of the nonlinearity of equation 3. $\widehat{F_{ST}}$ and $\widehat{R_{ST}}$ in the lower tail of their distribution will translate into a very large estimate of Nm, which will then take an unduly large weight in the average M estimates. The strong constraint in allele size imposed by Gaggiotti *et al*. (1999) might explain the remaining discrepancies between our simula-tions and theirs.

As noted above, it is under stepwise mutation that R_{ST} allow inference of quantities, such as the number of migrants, independently of mutation. If it is unlikely that the mutation scheme of microsatellite loci exactly follows this model, it remains open how much the mutation pattern of microsatellites deviates from it. An important concern is the effect of a finite number of possible allelic states. Constraints on allele size will render the mutation pattern more similar to that expected under a model where mutant alleles take one of k possible states at random (the 'KAM' model). Indeed, for the extreme case of a constraint to two allelic states, each mutation will generate the other allelic state. Rousset (1996) showed the expectations of identity in states under a KAM model to be those of an Infinite Allele Model (IAM), with a new mutation rate μ' modified according to the relation $\mu' = k\mu/(k-1)$. With constraint on allele size, size is no longer an accurate predictor of allelic distances. As the constraint gets stronger, alleles become equidistant. As perfect dinucleotide alleles rarely exceed 30 repetitions, the maximal possible size of microsatellite alleles seems to be constrained. The limited size of microsatellite alleles could be due to selection against long alleles (Garza *et al*. 1995; Nauta & Weissing 1996). It is however, more likely that the absence of very high repeat numbers is due to a mutation bias of long alleles towards shorter ones (Schlötterer *et al*. 1998). Xu *et al*. (2000) reported an excellent fit between empirical data and a mutation model where the rate of expansion mutations is constant across the entire allele distribution, and the rate of contraction mutations increases exponentially with allele size.

The accuracy of the stepwise mutation model to describe the evolution of microsatellite alleles remains an open question, particularly for the commonly used di-nucleotide repeat motifs (Ellegren 2000a, 2000b). However, our simulations show that even under the strictest stepwise conditions, differentiation statistics developed for this mutation model are not always the most adequate, because of their high associated variance. Departures from a generalized stepwise mutation model are likely to make the expectations of R_{ST} and F_{ST} statistics converge. As the estimators of the former will retain a larger variance than the latter, unless a prohibitive number of loci is used, $\widehat{F_{ST}}$ will often be the statistic of choice, particularly for moderate to high levels of gene flow.

Acknowledgements

We thank Laurent Excoffier, Pierre Fontanillas, Oscar Gaggiotti, Nicolas Lugon-Moulin, Marek Kimmel, Eric Petit, François Rousset and two anonymous reviewers for helpful discussions and comments on an earlier draft of this manuscript. F.B. was supported by the BBSRC. J.G. was supported by grant 98.055945 of the Swiss National Science Foundation, and by a matching fund subsidy from the University of Lausanne.

References

Balloux F (2001) EASYPOP Version 1.7: a computer program for population genetics simulations. *Journal of Heredity*, **92**, 301–302. [Available from http://www.unil.ch/izea/softwares/easypop.html].

Balloux F, Brünner H, Lugon-Moulin N, Hausser J, Goudet J (2000) Microsatellites can be misleading: an empirical and simulation study. *Evolution*, **54**, 1414–1422.

Brinkman B, Klintschar M, Neuhuber F, Hühne J, Burkhard F (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *American Journal of Human Genetics*, **62**, 1408–1415.

Chakraborty R, Nei M (1982) Genetic differentiation of quantitative characters between populations of species. I. Mutation and random genetic drift. *Genetical Research Cambridge*, **39**, 303–314.

Charlesworth B (1998) Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution*, **15**, 538–543.

Cockerham CC, Weir BS (1993) Estimation of gene flow from F-statistics. *Evolution*, **47**, 855–863.

Di Rienzo A, Donnely P, Toomajian C, *et al*. (1998) Heterogeneity of microsatellite mutations within and between loci, and implications for human demographic histories. *Genetics*, **148**, 1269–1284.

Ellegren H (2000a) Heterogeneous mutation process in human microsatellite DNA sequences. *Nature Genetics*, **24**, 400–402.

Ellegren H (2000b) Microsatellite mutations in the germline, implications for evolutionary inference. *Trends in Genetics*, **16**, 551–558.

Gaggiotti OE, Lange O, Rassman K, Gliddon C (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology*, **8**, 1513–1520.

Garza JC, Slatkin M, Freimer NF (1995) Microsatellite allele frequencies in humans and chimpanzees, with implications for constraints on allele size. *Molecular Biology and Evolution*, **12**, 594–603.

Goldstein DB, Linares AR, Feldman M, Cavalli-Sforza LL (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics*, **139**, 463–471.

Goodman SJ (1997) Rstcalc, a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molecular Ecology*, **6**, 881–885.

Goudet J (1995) FSTAT v1.2. A computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.

Goudet J (2001) *FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices* Version 2.9.3. Available from http://www.unil.ch/izea/softwares/fstat.html [updatedfrom Goudet (1995)].

Hedrick PW (1999) Highly variable loci and their interpretation in evolution and conservation. *Evolution*, **53**, 313–318.

Kimmel M, Chakraborty R (1996) Measures of variation at DNA repeat loci under a general stepwise mutation model. *Theoretical Population Biology*, **50**, 345–367.

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

King JP, Kimmel M, Chakraborty R (2000) A power analysis of microsatellite-based statistics for inferring past population growth. *Molecular Biology and Evolution*, **17**, 1859–1868.

Lugon-Moulin N, Brünner H, Wyttenbach A, Hausser J, Goudet J (1999) Hierarchical analysis of genetic differentiation in a hybrid zone of Sorex araneus (Insectivora, Soricidae). *Molecular Ecology*, **8**, 419–431.

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

Nagylaki T (1998) Fixation indices in subdivided populations. *Genetics*, **148**, 1325–1332.

Nauta MJ, Weissing FJ (1996) Constraints on allele size at microsatellite loci, implications for genetic differentiation. *Genetics*, **143**, 1021–1032.

Pritchard JK, Feldman MW (1996) Statistics for microsatellite variation based on the coalescence. *Theoretical Population Biology*, **50**, 325–344.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Rice JA (1995) *Mathematical Statistics and Data Analysis*, 2nd edn. Duxbury Press, Belmont CA.

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

Schlötterer C, Ritter R, Harr B, Brem G (1998) High mutation rate of a long microsatellite allele in Drosophila melanogaster provides evidence for allele-specific mutation rates. *Molecular Biology and Evolution*, **15**, 1269–1274.

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

Weber JL, Wong C (1993) Mutation of human short tandem repeats. *Human Molecular Genetics*, **2**, 1123–1128.

Wehrhahn CF (1975) The evolution of selectively similar electrophoretically detectable alleles in finite natural populations. *Genetics*, **80**, 375–394.

Weir BS (1996) *Genetic Data Analysis II*. Sinauer Associates, Sunderland, MA.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration. *Heredity*, **82**, 117–125.

Wolfram S (1991) *Mathematica, a System for Doing Mathematics by Computer*. Addison-Wesley Publishing Co, Redwood city, CA.

Wright S (1978) *Evolution and the Genetics of Population, Variability Within and Among Natural Populations*. University of Chicago Press, Chicago.

Xu X, Peng M, Fang Z, Xu X (2000) The direction of microsatellite mutations is dependent upon allele length. *Nature Genetics*, **24**, 396–399.

François Balloux is currently a postdoctoral fellow at the University of Edinburgh. His current research focuses on how population subdivision affects the evolutionary outcome in classical biological problems as the result of competition between sexually reproducing populations and asexual lineages or the evolution of virulence. Jérôme Goudet has been recently appointed professor in population genetics at the Institute of Ecology. His research focuses on theoretical and experimental aspects of mating systems in plants and animals, and more generally, on evolution in structured populations.

Appendix 1

Following the implementation of the island model in EASYPOP, we assume a diploid, monoecious population with nonoverlapping generation, where selfing occurs at random. Dispersal occurs at a rate m at the zygotic stage, and sampling takes place after dispersal, before reproduction and death. The population is made of a finite number n of demes, each of equal size N. Symmetric single-step mutation occurs at a rate μ. Under these conditions, the generating function of the probabilities p_k that a randomly chosen allele differs by k steps from another randomly chosen allele, $\psi_j(z) = \sum_{-\infty}^{+\infty} p_{k,j} z^k$, can be used to obtain the expectations (or theoretical values) for variance in allele sizes and gene diversity and therefore for R- and F-statistics (see Wehrhahn 1975; Rousset 1996). In the following, the generating functions $\psi_1(z)$ is for pairs of genes within individuals, $\psi_2(z)$ is for pairs of genes between individuals within demes and $\psi_3(z)$ is for pairs of genes between subpopulations. The recursions for the generating functions are

$$\left\{ \begin{aligned} \psi_{1,t+1}(z) &= r(z)\left[\frac{1}{N}\left(\frac{1+\psi_{1,t}(z)}{2}\right) + \left(1-\frac{1}{N}\right)\psi_{2,t}(z)\right] \\ \psi_{2,t+1}(z) &= r(z)\left[a\left(\frac{1}{N}\left(\frac{1+\psi_{1,t}(z)}{2}\right) + \left(1-\frac{1}{N}\right)\psi_{2,t}(z)\right)\right. \\ &\qquad \left. + (1-a)\psi_{3,t}(z)\right] \\ \psi_{3,t+1}(z) &= r(z)\left[b\left(\frac{1}{N}\left(\frac{1+\psi_{1,t}(z)}{2}\right) + \left(1-\frac{1}{N}\right)\psi_{2,t}(z)\right)\right. \\ &\qquad \left. + (1-b)\psi_{3,t}(z)\right] \end{aligned} \right.$$

where $a = (1-m)^2 + m^2/(n-1)$ is the proportion of pairs of genes within a deme which came from the same deme in the previous generation, $b = (1-a)/(n-1)$ is the proportion of pairs of genes in different demes which came from the same deme in the previous generation and $r(z) = (1-(2-z-1/z)u/2)^2$ is the factor by which stepwise mutation changes the generating functions (see Rousset 1996). At equilibrium between mutation, migration and drift, $\Psi_{j,t+1}(z) = \Psi_{j,t}(z) = \Psi_j(z)$ leading to the following system of equations:

$$\psi_1(z) = \frac{cr(z)[1+(b-1)r(z)]}{1+[b-1-c+a(2c-1)]r(z)+[a-b+c-2ac+bc]r(z)^2}$$

$$\psi_2(z) = \frac{cr(z)[a+(b-a)r(z)]}{1+[b-1-c+a(2c-1)]r(z)+[a-b+c-2ac+bc]r(z)^2}$$

$$\psi_3(z) = \frac{bcr(z)}{1+[b-1-c+a(2c-1)]r(z)+[a-b+c-2ac+bc]r(z)^2}$$

where $c = 1/2N$.

Since the expected mean difference in allele size with a symmetric mutation process is 0, the expected variances in allele size V_j is half the second derivative of the generating function of the probabilities p_k with respect to z in the neighbourhood of 1, $V_j = E[k_j^2]/2 = \frac{1}{2}[d^2\psi_j(z)/dz^2]\Big|_{z=1}$ (Rousset 1996). R_{ST} expectation is given by the ratio $(V_3 - V_2)/V_3$. Numerical expression for the V_j and R_{ST} were obtained with the computer package MATHEMATICA (Wolfram 1991).

For gene diversities, Rousset (1996) showed how to compute the probabilities of identity in state Q_1, Q_2 and Q_3. Using the numerical evaluation of the following expression:

$$Q_j = \frac{1}{\pi}\int_0^{\pi}\psi_j(e^{ix})dx$$

where the $\psi_j(z)$ have been defined above. The expectation of F_{ST} is then $(Q_2 - Q_3)/(1 - Q_3)$. Numerical expressions for the Q_j and F_{ST} were obtained with computer package MATHEMATICA (Wolfram 1991).

Appendix 2

Calling M_o and V_o the mean and variance in allele size at a locus, Goodman (1997) suggested that the centred normalized allele size $(x - M_o)/\sqrt{V_o}$ should be used instead of allele size x in the estimation of R_{ST}. Individual loci $\widehat{R_{\mathrm{ST}}}$ will not be affected by this centering–rescaling, but the overall loci estimator will, because each individual locus variance component will be divided by its locus allelic size variance [since $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$]. Goodman's $\widehat{R_{\mathrm{ST}}}$ will thus be expressed as:

$$\widehat{UR_{\mathrm{ST}}} = \frac{\sum_{i=1}^{l} \hat{V}_a^i / \hat{V}_o^i}{\sum_{i=1}^{l} \hat{V}_t^i / \hat{V}_o^i}$$

Where l represents the number of loci. V_t is an unbiased estimator of V_o, so that $\hat{V}_t \approx \hat{V}_o$, and therefore $\sum_{i=1}^{l} \hat{V}_t^i / \hat{V}_o^i \approx l$. (In fact, one could argue that V_t should be used for the total variance instead of V_o).

This leaves us with $\widehat{UR_{\mathrm{ST}}} \approx 1/l \sum_{i=1}^{l} \frac{\hat{V}_a^i}{\hat{V}_t^i} = \widehat{R_{\mathrm{STu}}}$ i.e. the average of the ratios. Therefore, the Goodman multilocus $\widehat{R_{\mathrm{ST}}}$ is similar to taking the average of individual loci $\widehat{R_{\mathrm{ST}}}$.