Tests for sex-biased dispersal using bi-parentally inherited genetic markers

JÉRÔME GOUDET,* NICOLAS PERRIN* and PETER WASER†
*Institute of Ecology, Biology Building, University of Lausanne, CH-1015 Lausanne, Switzerland, †Department of Biological Sciences, Purdue University, W. Lafayette IN 47907 USA

Abstract

Understanding why dispersal is sex-biased in many taxa is still a major concern in evolutionary ecology. Dispersal tends to be male-biased in mammals and female-biased in birds, but counter-examples exist and little is known about sex bias in other taxa. Obtaining accurate measures of dispersal in the field remains a problem. Here we describe and compare several methods for detecting sex-biased dispersal using bi-parentally inherited, codominant genetic markers. If gene flow is restricted among populations, then the genotype of an individual tells something about its origin. Provided that dispersal occurs at the juvenile stage and that sampling is carried out on adults, genotypes sampled from the dispersing sex should on average be less likely (compared to genotypes from the philopatric sex) in the population in which they were sampled. The dispersing sex should be less genetically structured and should present a larger heterozygote deficit. In this study we use computer simulations and a permutation test on four statistics to investigate the conditions under which sex-biased dispersal can be detected. Two tests emerge as fairly powerful. We present results concerning the optimal sampling strategy (varying number of samples, individuals, loci per individual and level of polymorphism) under different amounts of dispersal for each sex. These tests for biases in dispersal are also appropriate for any attribute (e.g. size, colour, status) suspected to influence the probability of dispersal. A windows program carrying out these tests can be freely downloaded from http://www.unil.ch/izea/softwares/fstat.html

Keywords: assignment index, *F*-statistics, individual-based models, microsatellites, population structure, randomizations

Received 22 November 2001; revision received 1 February 2002; accepted 1 February 2002

Introduction

Male-biased dispersal is pervasive in mammals, female-biased dispersal in birds (Greenwood 1980; Wolff 1994; Clarke *et al*. 1997). Numerous hypotheses have related dispersal sex biases to differences in the advantages that philopatry conveys to males and females in competition for mates or breeding resources, or to the impact of male–female differences on the fecundity costs of dispersal (Greenwood 1980; Dobson 1982; Johnson 1986; Pusey 1987; Part 1995; Zera & Denno 1997; Wolff & Plissner 1998). Recent models consider the consequences of dispersal for inbreeding and kin competition together, to generate qualitative predictions about male and female dispersal

between populations (Perrin & Mazalov 1999; 2000; Perrin & Goudet 2001). Ideas developed to explain dispersal in general can often be tested by their ability to explain differences in dispersal between the sexes.

As data on avian and mammalian dispersal have accumulated, it has become increasingly obvious that there are exceptions to the general patterns – mammals that show female-biased dispersal, or birds in which dispersers are primarily males. Species that depart from the 'norm' are fertile ground for exploring ideas about the evolution of dispersal tendencies (Greenwood 1980; Clutton-Brock 1989; Clark *et al*. 1997; Wolff & Plissner 1998). So are species whose dispersal patterns are unknown but whose mating systems or life history characteristics allow clear a priori predictions about dispersal sex bias. In this light, it might even be viewed as an advantage that we know virtually nothing about patterns of sex-biased dispersal in taxa other

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

than birds or mammals. The absence of data provides ample scope for making genuine predictions.

Data on the magnitude and direction of sex bias in dispersal are limited, because the well-known difficulties associated with estimating dispersal rates by following individual animals (e.g. Slatkin 1985; Baker *et al.* 1995; Koenig *et al.* 1996) apply with equal strength to measuring sex bias in those rates. The possibility of using 'indirect' methods to infer sex-specific dispersal rates from the spatial distribution of alleles or genotypes presents a seductive alternative.

There is some disagreement as to whether the indirect approach should be powerful enough to detect a sex bias in dispersal. For example, even if only one sex disperses, unusual alleles brought into the population by immigrants are transmitted, one generation later, to both sexes. It is therefore important to exclude 'predispersal' individuals when searching for a sex bias by comparing male and female genotypes. Also, if immigrants are rare, then their genotypes may have a small effect on the overall distribution of variation among males vs. females; while if immigrants are very common, then both male and female genotypes may be undifferentiated among populations. Thus comparisons of male and female genotypes may not have much statistical power under some conditions.

Nevertheless, even in the era of allozymes a few authors (McCracken 1984; Kawata 1985) used this approach. The advent of mini and microsatellites opened the possibility of much greater sensitivity in testing with genetic data whether sex-biased dispersal occurs (Favre *et al.* 1997; Ishibashi *et al.* 1997; Gompper *et al.* 1998; Knight *et al.* 1999; Mossman & Waser 1999; Stow *et al.* 2001). It is the aim of this study to investigate the conditions under which sex-biased dispersal can be detected with codominant, hypervariable markers such as microsatellites, comparing several alternative approaches. Two of these approaches are based on comparing measures of population structure for males and females using *F*-statistics; two are based on recently described individual 'assignment indices'.

Methods

Design of the tests

We assume a species with nonoverlapping generations where dispersal occurs at the juvenile stage, before reproduction. We assume further that individuals are sampled post dispersal. Under these conditions, several statistical descriptors of an individual's genotype can be used to indicate sex biases in dispersal. We explore the performance of four of them. The first two are traditional, 'global' descriptors of population structure, while the others are based on a more recent approach relying on individual genotypes.

1 F_{IS}: F_{IS} is a statistic describing how well the genotype frequencies within populations fit with Hardy–Weinberg expectation (Hartl & Clark 1997). If only males disperse, the males sampled from a single patch will be a mixture of two populations, residents and immigrants; due to the Wahlund effect, the sample should show a heterozygote deficit and a positive F_{IS}. In general, members of the dispersing sex should therefore display a higher F_{IS} than the more philopatric sex. Among the several estimators of F_{IS}, we chose Weir & Cockerham's (1984) because it is the most commonly used and it is also unbiased.

2 F_{ST}: F_{ST} is a statistic expressing the proportion of the total genetic variance that resides among populations (Hartl & Clark 1997). Allelic frequencies for individuals of the dispersing sex should be more similar than those for individuals of the more philopatric sex. We therefore expect F_{ST} for the more philopatric sex to be higher than that of the more dispersing sex. Among the available estimators of F_{ST}, we again chose Weir & Cockerham's (1984) for the same reasons as for F_{IS}. Relatedness has often been used to test for sex-bias differences (e.g. Ishibashi *et al.* 1997; Knight *et al.* 1999; Surridge *et al.* 1999). This statistic is connected to F_{ST} through the relation: $r = 2F_{ST}/(1 + F_{IT})$ (Hamilton 1971; Queller & Goodnight 1989). Since F_{IT} expectation is identical for the philopatric and the dispersing sex, a test based on relatedness is expected to perform essentially like a test based on F_{ST}. To verify this prediction, we also performed tests using relatedness as a statistic.

Assignment index AI: This statistic was first introduced in Paetkau *et al.* (1995) and modified later by Favre *et al.* (1997). For each individual j in locality k, we calculate the probability P_{kij} that its genotype at locus i should appear in the kth sample as the squared frequency of the allele if the individual is homozygous, or twice the product of the frequencies of its two alleles if it is heterozygous. In the absence of linkage disequilibrium, the probability of occurrence of a multilocus genotype is the product of the probabilities of the individual loci. Because populations can contain very different levels of gene diversity, the multilocus probabilities of individuals in different populations are not directly comparable. To remove this problem, the average probability of the sample is subtracted from the individual multilocus probability after log-transformation to avoid rounding errors with very small numbers. This gives the following formula for AIc, the corrected assignment index of individual j in sample k:

$$AIc_{kj} = Log\left[\prod_{i=1}^{l} P_{kij}\right] - 1/n\sum_{j=1}^{n}Log\left[\prod_{i=1}^{l}P_{kij}\right] \qquad (1)$$

for l loci and n individuals in the kth sample. The distribution of AIc will therefore be centred on 0. A positive value indicates a genotype more likely than average to

occur in its sample (probably a resident individual), while a negative value indicates a genotype less likely than average (potentially a disperser).

3 Mean of AIc *(mAIc):* Because immigrants tend to have lower *AIc* values than residents, under sex-biased dispersal we expect the sex that disperses most to have a lower *AIc* on average than the more philopatric sex.

4 Variance of AIc *(vAIc):* Because members of the dispersing sex will include both residents (with common genotypes) and immigrants (with rare genotypes), *vAIc* for the sex dispersing most should be largest.

To test whether these four statistics differ significantly between the two sexes, we used a randomization approach. Under the null hypothesis that males and females disperse equally, the four statistics do not depend on the variable 'sex'. Letting X_d and X_p be the statistic of interest for the dispersing and the philopatric sex, respectively, we proceeded as follows: (i) We first calculated the statistic for each sex over all populations and either take the difference $(X_d^\circ - X_p^\circ)$ for F_{IS}, the difference $(X_p^\circ - X_d^\circ)$ for *mAIc* and F_{ST}; or the ratio X_d° / X_p° for *vAIc*; (ii) We assigned a sex randomly to each of the multilocus genotypes (keeping the genotypes in their original sample, and the sex ratio in each sample constant); (iii) We recalculated the appropriate difference or ratio for the randomized data set; (iv) We repeated steps (i–iii) 100 times. For the test based on the variance of *AIc*, a ratio rather than a difference was used because statistical tests designed to compare variances are usually ratios (for instance the ANOVA *F*-test).

The probability that dispersal is unbiased by sex was then estimated as the proportion of times where the relevant statistic obtained from the randomized data sets is larger or equal to the statistic obtained from the observed data set. Power for each test was estimated as the proportion of simulated data sets giving a significant result at the 0.05 nominal level. A test was considered powerful when 80% of the simulated data sets gave a significant result. The tests just described are one-tailed, because we had an a priori idea of which sex is likely to disperse most. Two-tailed tests could also be constructed under the same principle using either the absolute value of the differences, or the ratio of the largest to smallest variance. These tests were expected to have reduced power.

Individual-based Monte Carlo simulations

In order to estimate the power of these tests under different scenarios, we used the individual-based Monte Carlo computer simulation programs GENETIX and EASYPOP (available from the web at http://bilbo.bio.purdue.edu/~pwaser/genetix and http://www.unil.ch/izea/softwares/easypop.html (Balloux 2001); these two programs were used for cross-

validation of the simulations). The main parameters of the simulations are the number of populations (p), the number of males and females per population ($N_m = N_f$) and the male and female dispersal rates d_m and d_f. In each generation, reproduction is followed by juvenile dispersal (and adult death), then by sampling. Dispersal follows an island model, so that each population receives on average proportions d_m and d_f of dispersing males and females coming from the other populations, and sends proportions $d_m/(p-1)$ and $d_f/(p-1)$ of males and females born locally to each of the other populations. Each population is made of an equal number of males and females, who reproduce by mating at random. Each individual is characterized by a sex locus and l autosomal loci, genotypes of which are determined by Mendelian segregation of parental alleles. For autosomal loci, mutation occurs at a rate μ during gametogenesis, each new mutation giving rise to a new state chosen among K possible (the K allele model of mutation, KAM). We ran each simulation for 1000 generations, which was enough for genetic diversity to reach equilibrium between mutation, migration and drift.

Parameters for the simulations

Most simulations were run with $p = 24$, $N_m = N_f = 12$ and l = 8. The mutation rate was set to 0.001, implying a level of genetic diversity close to 0.7, a value commonly found with microsatellite markers (Jarne & Lagoda 1996); the number of possible allelic states was fixed at 25, a number corresponding to the range in size of many published microsatellite markers (Goldstein & Pollock 1997). The following per capita dispersal rates were analysed: no dispersal, 1%, 2%, 4%, 8%, 16%, 32% and finally 50%. For each dispersal rate, we also varied the intensity of the sex bias by varying the sex ratio of dispersers from 100:0, 90:10, … 50:50. Simulations with no dispersal, or with a 50:50 sex ratio of dispersers, satisfy the null hypothesis of no sex-biased dispersal. For most of these parameter sets, we generated 1000 replicates.

We also ran a restricted set of simulations to investigate the effects of the number of loci and of population size. For the number of loci, we ran 100 simulations with 50 loci instead of eight. Dispersal rate was fixed at 16% with a 70:30 bias, corresponding to a situation where the power was below 80% for the different tests. To investigate the effect of population size, we ran simulations with $p = 40$ and $N_m = N_f = 50$. Dispersal rates were set to 50%, 32%, 8% and 2%, and bias was fixed at 100:0. At dispersal rate 32% we also looked at an 80:20 bias.

Effects of sampling strategies, levels of polymorphism and number of loci

In most field situations, it is not possible to sample the whole population. Therefore, in addition to the results for

an exhaustive sampling scheme, we report for the main set of simulations the power of the different tests when sampling only 25% of the individuals under three sampling strategies: 12 individuals in 12 populations, six individuals in 24 populations and 24 individuals in six populations. We also report results for the restricted set of larger simulations in which we subsampled either 24 individuals in 24 populations, or 100 individuals in six populations.

Two other factors likely to affect the power of the tests are the level of polymorphism of the markers and the number of loci available for analysis. The level of polymorphism of the markers was varied by pooling alleles into two classes (e.g. allozymes) rather than 25. As for the number of loci, we analyse the simulations' results using only a subset of the markers (2, 4, 8, 20 or 50).

To illustrate how these tests work, we used the published data set of Favre *et al.* (1997) and Balloux *et al.* (1998), available from http://www.unil.ch/izea/datasets.html. The data set consists of 8–12 sexed individuals from 15 localities typed at eight microsatellites. Favre *et al.* (1997) showed that the *AIc* of females had a significantly lower median and a significantly larger variance than that of males, and observed that the distribution of *AIc* in females was bimodal. We apply the four tests described to this data set, using 1000 randomizations to generate the distribution

of the different statistics under the null hypothesis of no bias in dispersal.

Results

As expected, all the tests have a power close to 5% when the null hypothesis is true. The relatedness-based test is very similar in power to the test based on F_{ST}. Therefore, only the result of the latter will be shown in detail. Because the power of the test based on F_{IS} never reached 70%, the results of this test will not be presented in detail.

Exhaustive sampling

Figure 1 displays the power of the other three tests with exhaustive sampling (24 individuals in 24 populations typed at eight highly variable loci). When bias is maximal (100:0; panel a), the three tests have a high power (larger than 80% in most situations). When dispersal rate is very low (< = 2%), the *vAIc* test is the most powerful (and the only test with a power over 80%), followed by the *mAIc* test and the F_{ST} test. At higher dispersal rates (> = 4%), the power of *mAIc* and F_{ST} tests increases and quickly reaches the 80% threshold, while that of the *vAIc* test drops, and falls below 80% for dispersal rates above 30%.
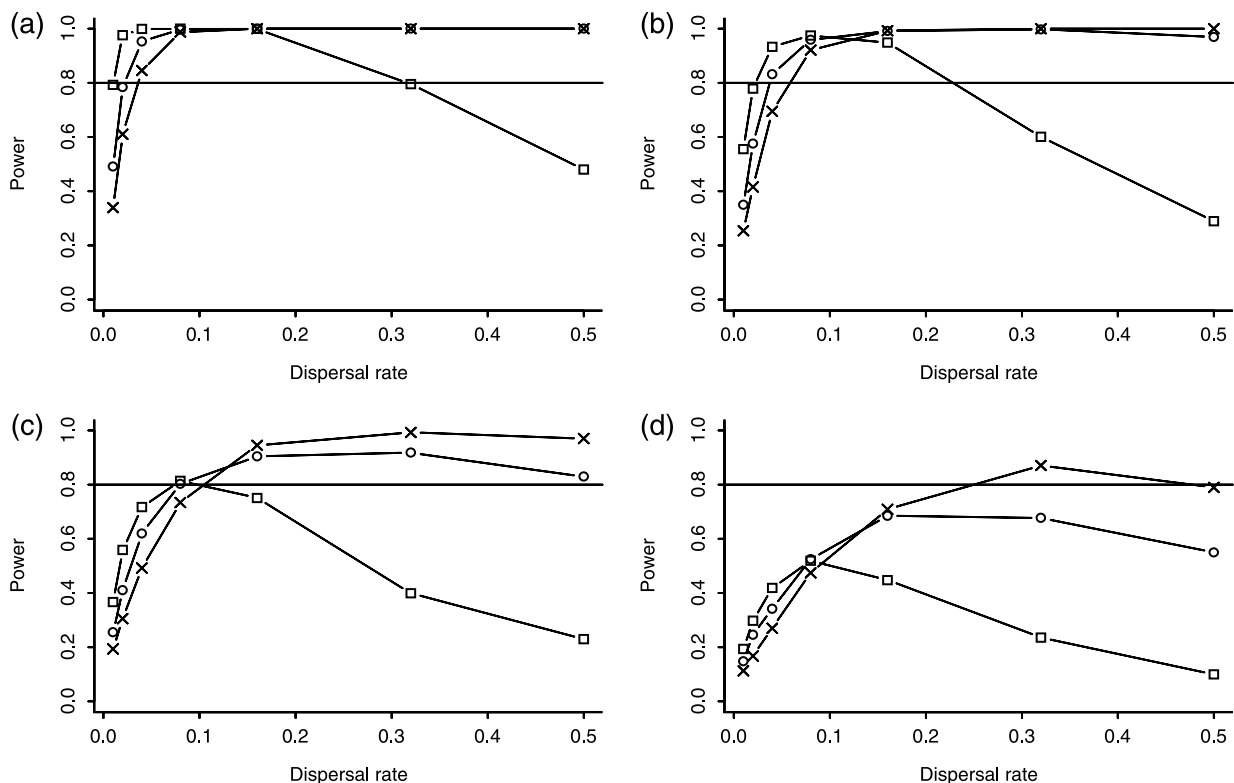

Fig. 1 Effect of the dispersal rate for different biases in dispersal on the power of the *vAIc*, the *mAIc* and the F_{ST} tests for an exhaustive sampling of the population when individuals are typed at eight highly variable loci. Square: *vAIc* test; circle: *mAIc* test; cross: F_{ST} test. Panel a: 100:0 bias in dispersal; panel b: 90:10; panel c: 80:20; panel d: 70:30.
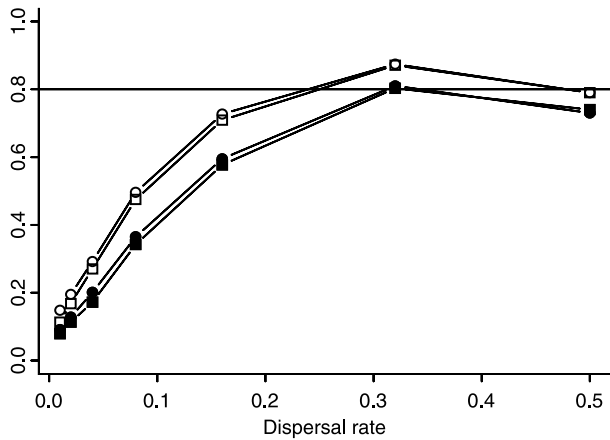
Fig. 2 Similarity of the power of tests based on F_{ST} (squares) and relatedness (circles), for one-sided (open symbols) and two-sided (filled symbols) tests. Open symbols: one-sided tests. Filled symbols: two-sided tests. Circles: relatedness test. Square: F_{ST} tests. The results shown are for exhaustive sampling, with a 70:30 sex bias in dispersal, for varying levels of dispersal.

In panels b and c, sex bias in dispersal is 90:10 and 80:20, respectively. F_{ST} and, to a lesser extent, *mAIc* tests remain powerful over most of the range of dispersal rates, while the *vAIc* test quickly loses power as dispersal rate increases.
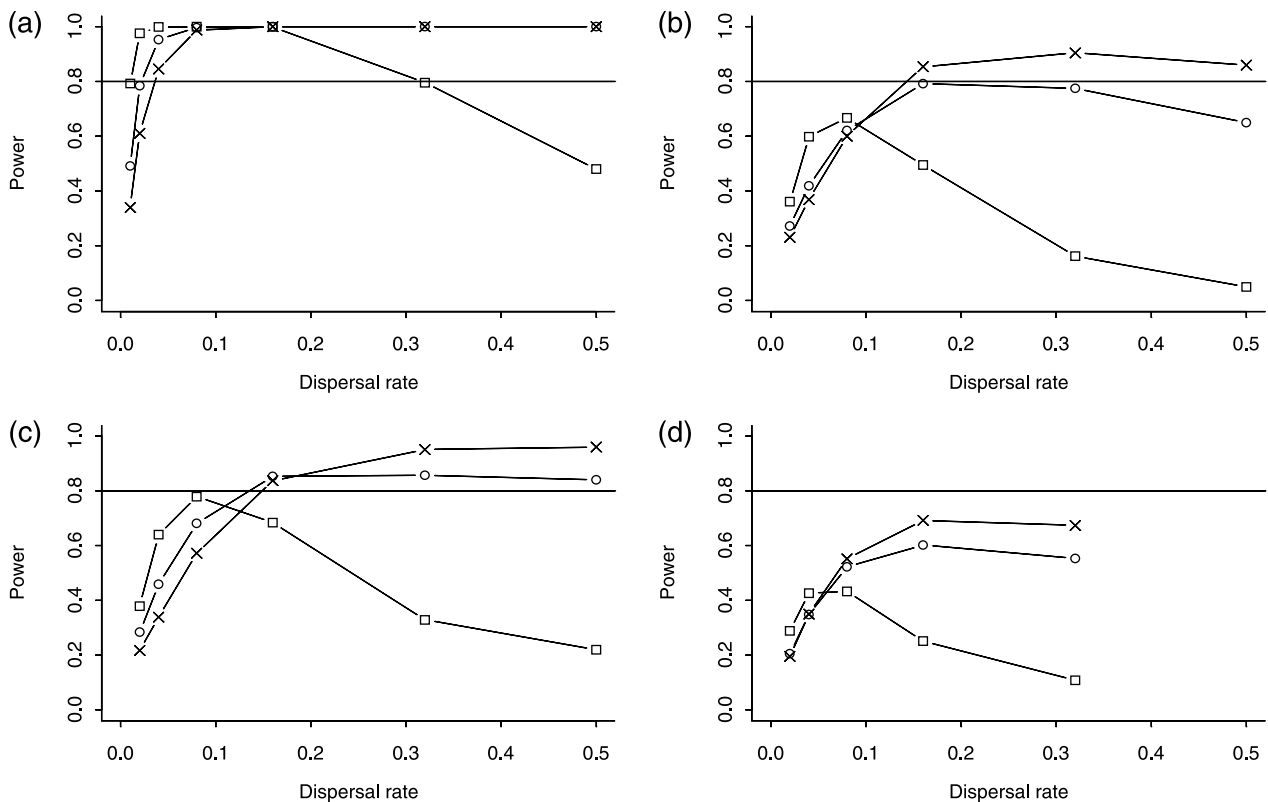
As asymmetry in dispersal decreases, so too does the power of the different tests. The power of tests to detect a 70:30 bias in dispersal (panel d, males still disperse more than twice as much as females) is very low. The *vAIc* test is most adversely affected followed by the *mAIc* test. The F_{ST} test barely reaches the 80% power threshold for a 32% dispersal rate.

Figure 2 illustrates the similarity in power of the tests based on F_{ST} and relatedness. The average difference in power between these two tests over all the simulated scenarios is 1.4%, with the largest differences never exceeding 8% for dispersal rates of 1% and only one sex dispersing. The difference in power between these two tests disappears totally for dispersal rates larger than 16% and bias in dispersal lower than 80:20.

Figure 2 also shows that two-sided tests have a slightly reduced power compared with their one-sided equivalent (6% reduction on average, never exceeding a 15% reduction in power), as expected.

Partial sampling

The effect of the sampling strategy on the power of the different tests when dispersal is fully biased is displayed on Fig. 3. Panel a shows the tests' power when all the 576
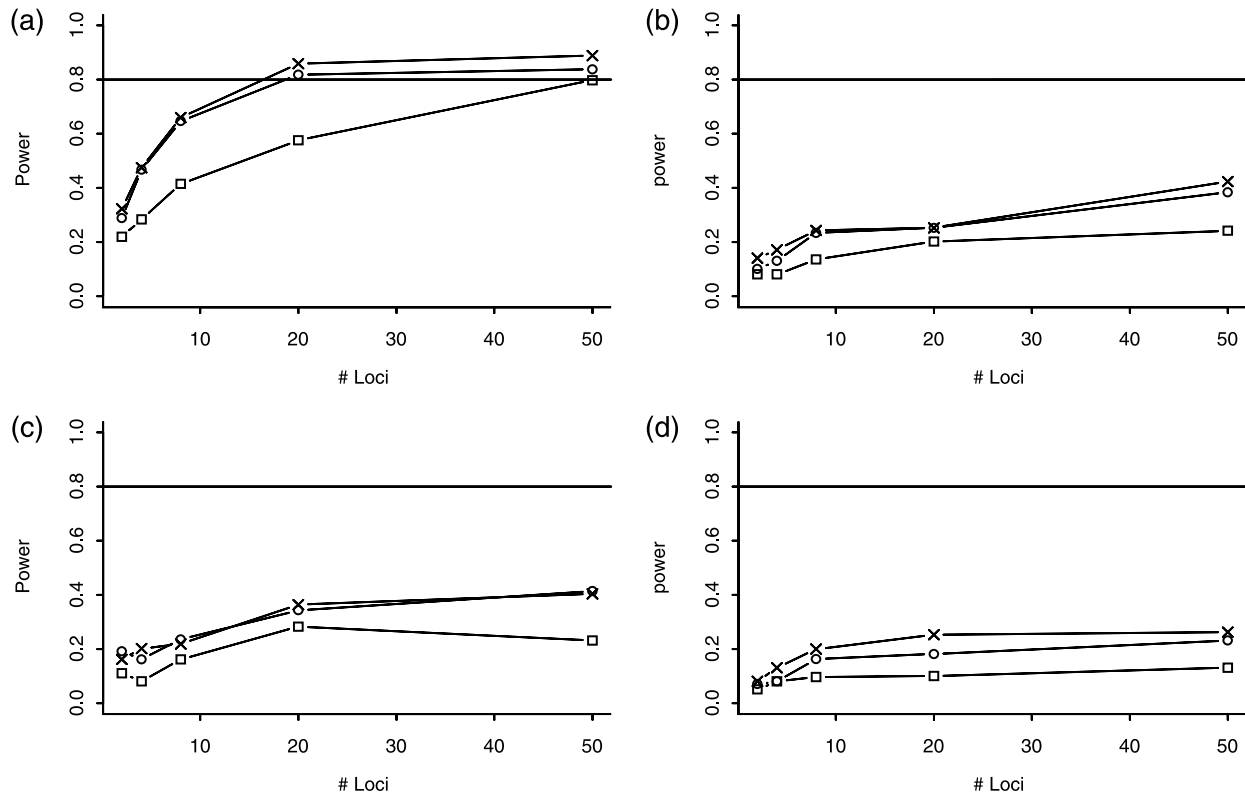

Fig. 3 Effect of dispersal rate for different sampling designs on the power of the *vAIc*, the *mAIc* and the F_{ST} tests when dispersal is 100:0 sex-biased. Individuals have been typed at eight highly variable loci. Square: *vAIc* test; circle: *mAIc* test; cross: F_{ST} test. (a) Exhaustive sampling (24 individuals in 24 populations); (b) 12 individuals in 12 populations; (c) 24 individuals in six populations; (d) six individuals in 24 populations.

Fig. 4 Effect of the number of typed loci for different sampling designs on the power of the *vAIc*, the *mAIc* and the F_{ST} tests for simulations where the sex bias is 70:30 and dispersal is 16%. The loci are highly variable. Square: *vAIc* test; circle: *mAIc* test; cross: F_{ST} test. (a) Exhaustive sampling (24 individuals in 24 populations); (b) 12 individuals in 12 populations; (c) 24 individuals in six populations; (d) six individuals in 24 populations.

individuals (24 animals in each of 24 populations) are sampled, while panels b, c and d display the results when only 144 individuals are sampled, under three different sampling schemes. Power of the tests is affected strongly by partial sampling. The *vAIc* test (which was the most powerful with low dispersal and complete sampling) never reaches the 80% threshold. The *mAIc* test barely reaches it when 24 individuals and six populations are sampled (panel c). In contrast, the F_{ST} test has reasonable power when dispersal rate exceeds 10%, and when either 12 individuals in 12 populations (panel b) or 24 individuals in six populations (panel c) are sampled. The worst sampling strategy is a few individuals in many populations (e.g. six individuals in 24 populations, panel d), where none of the tests reaches the 80% threshold.

When the sex bias in dispersal is not complete, the power for the different tests drops even more. For instance, with a 80:20 bias and sampling 144 individuals, none of the tests reached the 80% threshold under any of the three sampling strategies (data not shown). With a 90:10 bias, only the F_{ST} test reaches the 80% threshold, and only for a dispersal rate of 32% and exhaustive sampling of six populations.

Number of loci

For a 70:30 bias, 16% dispersal and eight loci, none of the tests had a power above 80% (see Fig. 1, panel d). Figure 4 shows the effect of increasing the number of typed loci. Panel (a) corresponds to exhaustive sampling, while panels b, c and d represent different combinations of partial sampling (12×12, 6×24 and 24×6, respectively). In the case of exhaustive sampling (panel a), adding loci increases the power of all tests. With 20 loci, tests based on *mAIc* and F_{ST} pass the 80% threshold. Further increasing the number of loci to 50 has little effect on these two tests but allows the power of *vAIc* test to reach 80%. All the tests have low power when sampling fewer than eight loci (panel a). When sampling is not exhaustive (panels b, c and d) none of the tests reach 80% power, no matter how many loci are used.

Level of polymorphism

As the level of polymorphism of the markers decreases, so does the power of the different tests (Fig. 5). When the number of alleles is reduced from 25 to two alleles (which
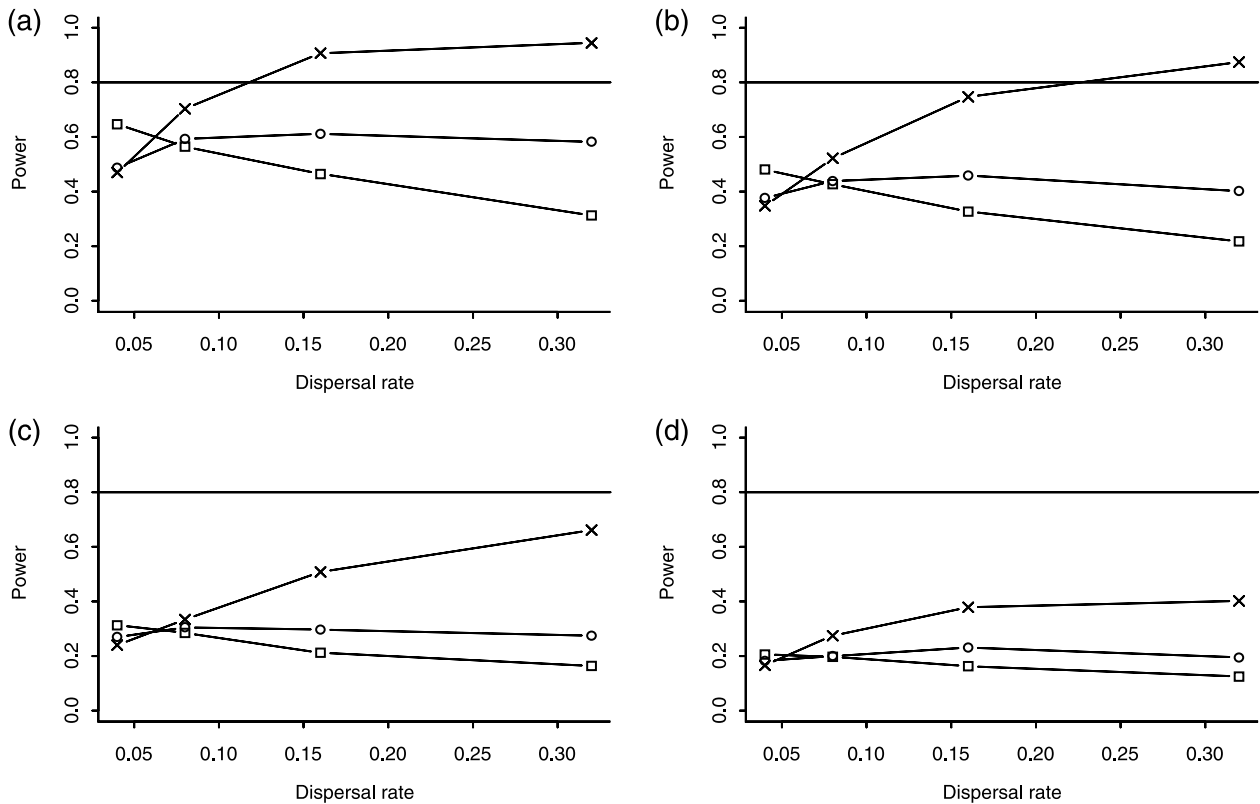
Fig. 5 As Fig. 1, but with only two alleles per locus and eight loci. Square: *vAIc* test; circle: *mAIc* test; cross: F_{ST} test. (a) 100:0 bias in dispersal; (b) 90:10; (c) 80:20; (d) 100:0 bias in dispersal, 12 individuals sampled in 12 populations.

halves the level of genetic diversity), the power is strongly reduced. With exhaustive sampling, F_{ST} reaches 80% power for the strong biases (100:0 and 90:10; panels a and b). With lower biases (panel c) or partial sampling (panel d), no tests reach 80% power.

Larger populations

The results for the larger populations are presented in Fig. 6. When sampling is exhaustive (40 populations of 100 individuals typed at 20 loci) and the bias maximal (100:0), power is generally very good (panel a). Apart from the *vAIc* test with 50% dispersal, all tests reach the 80% power threshold, and in most cases each replicate gives a significant result.

Sampling the population affects the results of the test only partially: while the power changes little when the number of populations sampled is dropped from 40 to four (compare panels a and b and d and e of Fig. 6), dropping the number of individuals sampled from 100 to 50 lowers it considerably (data not shown). When only 24 individuals are sampled in 24 populations (Fig. 6, panel c), F_{ST} retains a power close to or larger than 80% at high dispersal rates, *vAIc* reaches the threshold at low dispersal rates, but *mAIc* never attains it.

A less-than-complete sex bias can still be detected with the F_{ST} test, but not with the others, and only as long as at least some populations are sampled exhaustively (Fig. 6, compare panel b with panel c).

Dropping the number of typed loci from 20 to eight (Fig. 6, panels d, e and f) has a limited impact when sampling is exhaustive (panel d). When it is not (panels e and f), only the F_{ST} test remains powerful when bias in dispersal is complete and some populations are sampled extensively (panel e). Overall, the F_{ST}-based test seems to perform best, and to be the least sensitive to changes is the sampling scheme.

Application

Balloux *et al.* (1998) estimated from the *Crocidura russula* data set that about 40% of the females, and practically no males, disperse among sites within villages. As the trapping intensity was high, most individuals from these sites were probably captured. This data set therefore represents a situation close to ideal for detecting sex bias in dispersal. The results for the *mAIc*, *vAIc*, F_{IS} and F_{ST} tests are shown in Fig. 7. The *mAIc*, *vAIc* and F_{ST} tests are all highly significant, with no randomized data sets giving a test statistic larger than the observed
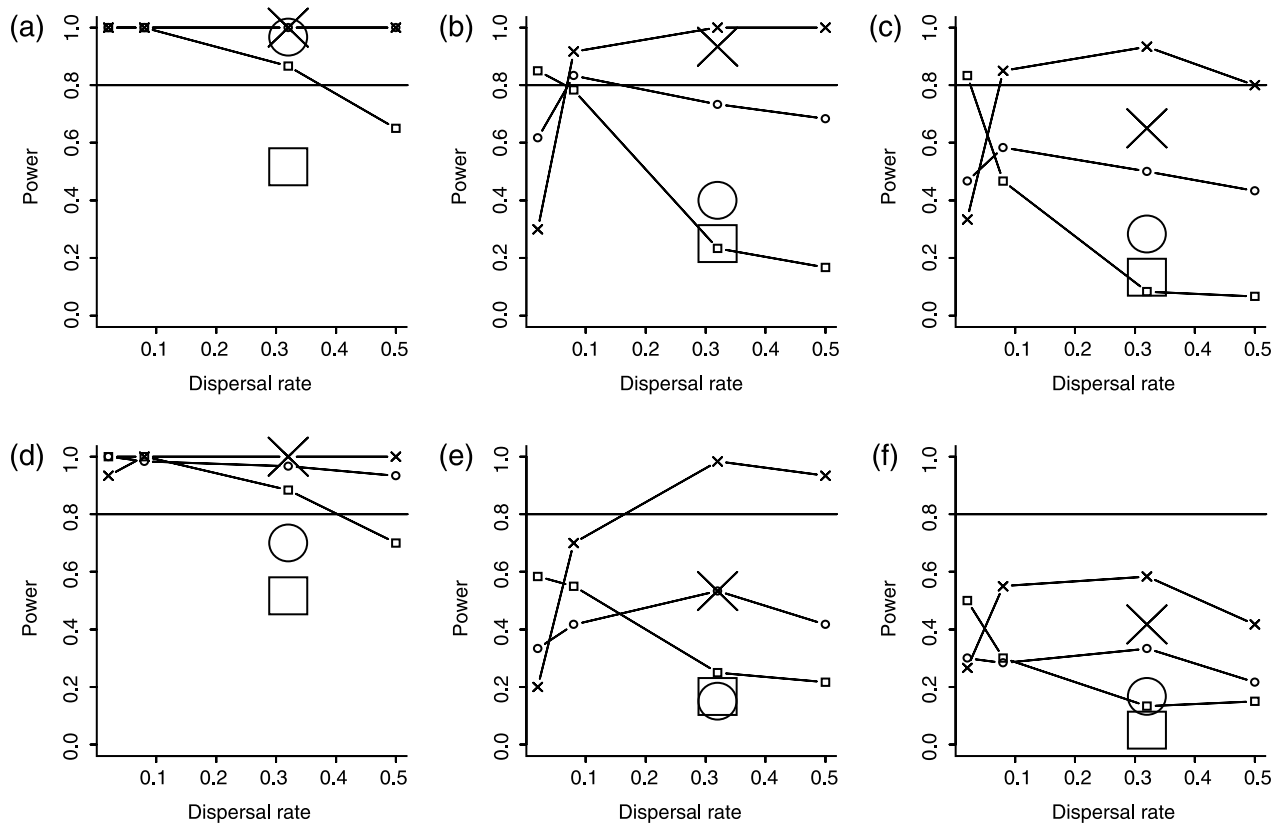
Fig. 6 Effect of dispersal rate for different sampling designs and number of typed loci on the power of the *vAIc*, the *mAIc* and the F_{ST} tests for the large simulations (100 individuals in 40 populations). Small symbols: dispersal is 100:0 sexbiased. Large symbols: dispersal is 80:20 sex-biased. Square: *vAIc* test; circle: *mAIc* test; cross: F_{ST} test. (a, b, c) 20 highly polymorphic loci; (d, e, f) eight highly polymorphic loci. (a, d) Exhaustive sampling. (b, e) 100 individuals in six populations. (c, f) 24 individuals in 24 populations.

($P < 0.001$ for all three tests). On the other hand, the F_{IS}-based test is not significant, with a P-value of 0.18.

Discussion

Under which conditions can a sex bias in dispersal be detected? From our results, many factors are likely to matter, including dispersal rate, bias intensity and sampling scheme.

Dispersal rate

In all simulations, power turns out to be a nonmonotonic function of dispersal rate, with maximal power for intermediate dispersal values (Fig. 1). This is obvious for the *vAIc* test, but also true for the F_{ST} and *mAIc* tests (e.g. Fig. 1 panel d; Fig. 6 panels c, e, f). This effect stems from the two contrasting consequences of dispersal: on one hand, when dispersal is very rare, immigrants constitute only a small proportion of the individuals sampled and may thus pass undetected. On the other hand, when dispersal is widespread, populations are less differentiated, so that the genotypes of immigrants are much less distinct. There is a clear trade-off between these two

aspects of dispersal, and optimal detection will occur for intermediate dispersal rates.

Interestingly, the dispersal rate at which sex bias is most readily detected differs among indices. *vAIc* performs best at low dispersal rate (less than 10%), while F_{ST} performs best at higher dispersal (more than 10%). Variance in *AIc* thus appears very sensitive to the presence of a few rare immigrants with very distinct genotypes, while F_{ST} is more sensitive if dispersers constitute a larger proportion of the sample, even if their genotypes are less distinct. *mAIc* is intermediate between these two tests.

Bias intensity

Sex bias needs to be very intense to be detected by any of these methods. If only one sex disperses (100:0), then bias can be detected most of the time provided sampling is exhaustive. However, bias is already significantly more difficult to detect as soon as its intensity drops to 80:20 (compare panels a and c, Fig. 1). In that case, only F_{ST} and *mAIc* have a chance of detecting a sex bias, and only when the dispersal rate is relatively high. There is little hope of detecting sex bias with any of these methods at intensities less than 80:20.
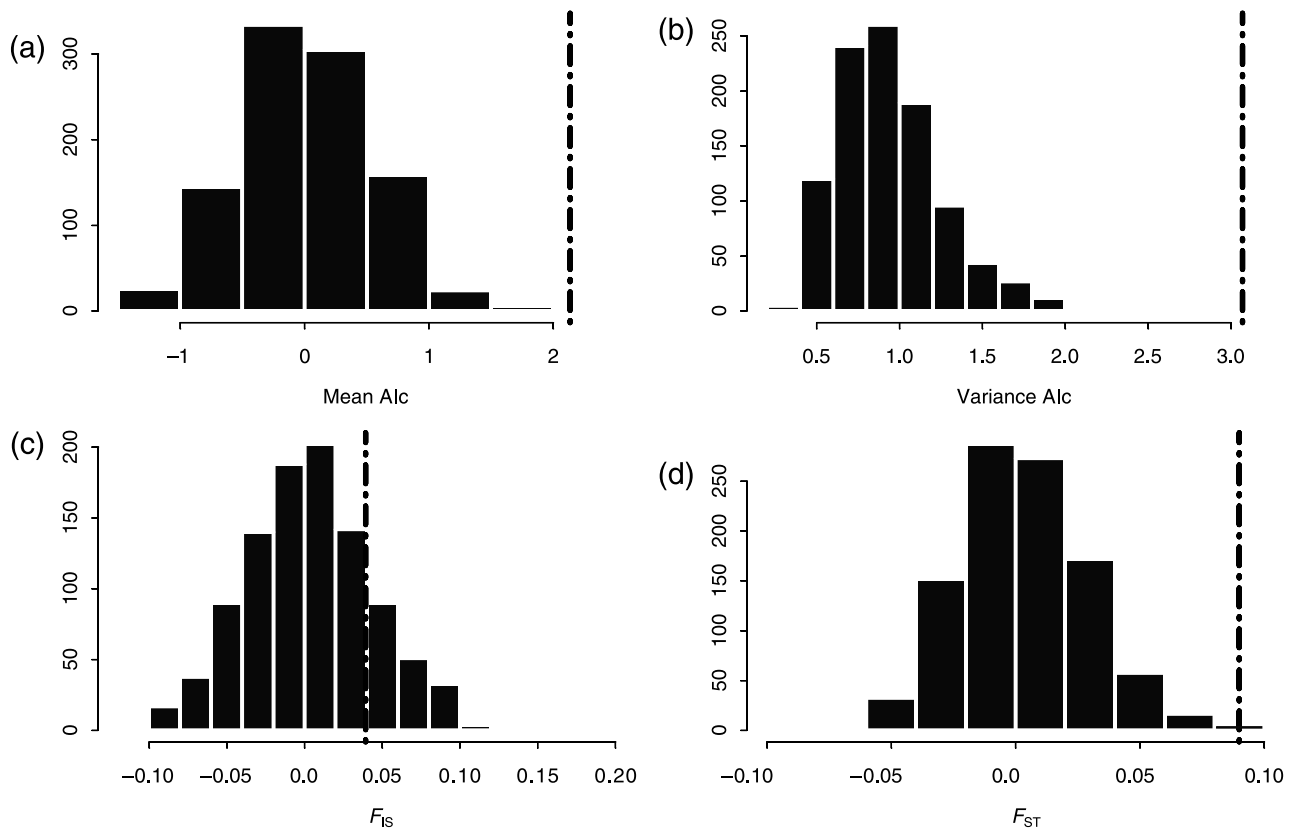
Fig. 7 Distribution of the randomized statistics for a real data set. The species is the greater white-toothed shrew, *Crocidura russula*. Eight microsatellites were typed from eight to 12 individuals in each of 15 samples. The vertical bar corresponds to the observed value of the test statistics. (a) *Mean AIc* test (*mean AIc* males/*mean AIc* females); (b) *variance AIc* test (*variance AIc* females/*variance AIc* males); (c) F_{IS} test (F_{IS} females/F_{IS} males); (d) F_{ST} test (F_{ST} males/F_{ST} females). *mAIc*, *vAIc* and F_{ST} tests are highly significant ($P < 0.001$) while F_{IS} test is not ($P = 0.18$).

Sampling design

Partial sampling strongly reduces the ability to demonstrate a bias (Figs 3 and 6). This implies that large sampling efforts are essential to detect sex-biased dispersal. Our results show that it is better to increase the number of individuals per population, rather than the number of populations sampled. An intuitive interpretation is that, with more individuals per sample, allele frequencies are better estimated, making it easier to distinguish between resident and immigrant genotypes. Note also that increasing sample size increases the likelihood that dispersing individuals will be present in the sample. This is obviously a prerequisite for any of the tests to work including those based on male–female differences in F_{ST} and F_{IS}. Unlike most uses of *F*-statistics, our male–female comparisons are quantifications of current, not past dispersal (because the genes of immigrants will be carried by both their daughters and sons, any difference between the sexes will disappear as soon as immigrants have reproduced). We note also that the rate of dispersal, rather than the absolute number of immigrants, is the crucial

parameter for these tests. It implies that the important factor for these tests is not F_{ST}, which depends on the number of dispersing individuals. Tests for biased dispersal rely on the proportion of dispersing individuals in the sample, which needs to be neither too small nor so large that samples appear undifferentiated.

Number of loci/alleles

When power is low due to a low bias in dispersal, increasing the number of loci might appear to be a good strategy. However, this strategy appears useful mainly when sampling design is exhaustive (many individuals, many populations, Fig. 4). Otherwise, we suggest investing in the number of individuals rather than the number of loci. Furthermore, if the investigator has access to a limited number of loci (e.g. < 20), then only F_{ST} and *mAIc* really gain power. They are even situations where increasing the number of loci for the *vAIc* test actually decreases power (see Fig. 4, panel c for instance).

As for the number of alleles, when it is small, as is generally the case with allozyme data, bias is detectable only

when it is intense (at least 90:10), and the only useful index is F_{ST} (Fig. 5).

Spatial pattern of dispersal

Are our results likely to be sensitive to our use of the island model? One way to answer this is to examine the opposite extreme model, a stepping-stone. In this model, the immigrants will be coming only from neighbouring populations. For the statistic F_{ST}, we expect to see differences between males and females for neighbouring populations, but these differences will vanish as populations are further and further apart. As a global F_{ST} statistic for each sex is an average of the pairwise statistics, it will therefore be an average of situations where there is a difference and situations where there is not. This is likely to make the F_{ST}-based test less powerful under isolation by distance. This lowering in power would be worsened if sampling is not exhaustive and the populations sampled are far apart. On the other hand, Aic-based tests should not be affected by a pattern of isolation by distance, as for these tests classes of individuals are compared within samples rather than among.

This is not necessarily a shortcoming of the F_{ST} test, however, as short-distance sex-biased dispersal might be easier to detect than long distance, because the proportion of short-distance dispersers is probably larger. Isolation by distance can be estimated by the regression of pairwise relatedness on pairwise geographical distance. The slope of the regression has been shown to differ among sexes in a few studies (e.g. Ishibashi *et al.* 1997; Knight *et al.* 1999). However, the scale at which sex bias dispersal is occurring might often be one where no isolation by distance can be evidenced. For instance, Balloux *et al.* (1998) found evidence of sex-biased dispersal by *C. russula* among localities within a village, a scale at which no isolation by distance could be shown. On the other hand, no sex bias in dispersal could be shown between villages, where isolation by distance was demonstrated. Sex biases in dispersal are likely to be aimed at avoiding inbreeding and/or kin-competition (Perrin & Goudet 2001), which are easily escaped by short-distance dispersal. By contrast, long-distance dispersal is more likely to be aimed at colonizing empty patches, in which case dispersal is unlikely to be sex-specific. Ji *et al.* (2001), for instance, show that brushtail possums (*Trichosurus vulpecula*) of both sexes disperse to colonize empty sites, while dispersal within occupied areas is male-biased.

Other indices of interest?

The test based on comparison of male and female F_{ST} generally performed well, except at low dispersal rates. Poor performance at low dispersal rates may be due to the estimator chosen. We used Weir & Cockerham's (1984) estimator of F_{ST}, which is unbiased but has a large variance (e.g. Goudet *et al.* 1996). Perhaps tests based on other estimators, such as Robertson & Hill's (1984), which is slightly biased but with a lower variance, would be more powerful at low dispersal rates.

We mentioned that relatedness could be used as a test statistic instead of F_{ST}. Figure 2 shows that tests based on these two statistics are indeed extremely close in power. Because band-sharing index can be used as a surrogate of relatedness (e.g. Gompper *et al.* 1998; Ji *et al.* 2001), our approach can be generalized easily to markers other than microsatellites (for instance fingerprints and AFLPs).

For *mAIc*, we used a crude test statistic, namely the difference between two means. For samples of small sizes, a test based on the pivotal *t*-statistic defined as

$$t = \frac{mAI_c^p - mAI_c^d}{\sqrt{Var(AI_c^p)/n_p + Var(AI_c^d)/n_d}}$$

(where n_p and n_d are the number of individuals in the more philopatric and the more dispersing group, respectively) have a slightly improved power (Manly 1997). However, the increase in power is very marginal (less than 5%), and only present when taking samples of very limited size (data not shown).

In considering the use of F_{ST} vs. *Aic*-based approaches, it is important to bear in mind that *AIc* statistics can be used for other purposes, such as the identification of dispersing genotypes or the assignment of individuals to populations (e.g. Rannala & Mountain 1997; Cornuet *et al.* 1999; Pritchard *et al.* 2000). Note also that if only one population has been sampled F_{ST} cannot be used, whereas *mAIc* and *vAIc* (and F_{IS}) can.

It has been pointed out frequently that differences in population structure as inferred from mitochondrial and nuclear markers imply sex differences in dispersal (e.g. Melnick & Hoelzer 1992; Rassmann *et al.* 1997; Paetkau *et al.* 1998; Gibbs *et al.* 2000; Petit *et al.* 2001). One might do even better by quantifying the differences in structuring measured with polymorphic male- and female-specific markers. In mammals, for instance, one could compare the F_{ST} values obtained from the mitochondrial D-loop and from Y chromosome microsatellites (Balloux *et al.* 2000; Kittles *et al.* 1999; Kalaydjieva *et al.* 2001). Note, however, that this approach would measure something different than the tests we report here; F_{ST} comparisons based on sex-specific genetic markers would reflect long-term effective gene flow rather than instantaneous dispersal rates.

Applicability of the tests

In which species are these tests likely to be applicable? Greenwood (1984) showed that sex bias is pervasive in birds and mammals; but sex-biased dispersal also exists in

cold-blooded vertebrates and invertebrates. For instance, sex-biased dispersal has been shown in lizards (Doughty *et al.* 1994; Stow *et al.* 2001), fish (Knight *et al.* 1999) and ants (Ross & Keller 1995). For the tests to be applicable, a first prerequisite is that dispersal occurs at the juvenile stage, before reproduction. A second prerequisite is that sampling occurs after the juveniles have dispersed. It may be difficult to obtain the adequate samples in ants, for instance, where male lifespan is short (Hölldobler & Wilson 1990) or in the noctule bat, where males do not join roosting sites and therefore escape capture (Petit *et al.* 2001). The tests will be applied more easily to species where the breeding groups are identified easily and of small size. Even extreme sex-bias in dispersal might go undetected if the proportion of immigrants in the sample is not sufficient.

While we focus on sex-biased differences in dispersal, an obvious generalization can be made for any individual status. For instance, we could test whether larger animals disperse more than smaller ones, whether parasitized individuals are less philopatric than healthy ones or whether inbred individuals are more vagile than outcrossed. For these tests to be valid, however, one needs to ensure that sampling occurs after dispersal and that the individuals sampled belong to the same cohort.

Conclusion

The tests presented have limited power unless the bias in dispersal is extreme, larger than 80:20. What matters is the proportion of dispersing individuals in the sample, not the overall level of differentiation, unless differentiation is so low that no sex differences in allele frequencies between samples can be detected. When dispersal is very low (lower than 10%), *vAIc* performs best. In all other cases, the statistic that is most powerful and least sensitive to changes in sampling scheme and magnitude of the sex-bias in dispersal is F_{ST}. It is better to sample more individuals per sample than many samples with few individuals. All these tests seem to depend strongly on the proportion of the population sampled, and gains in power are likely to be achieved by increasing the sampling intensity. The number of markers used to type individuals seems to matter little above eight. A Windows program carrying out these tests can be downloaded freely from http://www.unil.ch/izea/softwares/fstat.html

Acknowledgements

References

Baker M, Nur N, Geupel GR (1995) Correcting biased estimates of dispersal and survival due to limited study area: theory and an application using wrentits. *Condor*, **97**, 663–674.

Balloux F (2001) EASYPOP version 1.7: a computer program for population genetics simulations. *Journal of Heredity*, **92**, 301–302.

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

Balloux F, Goudet J, Perrin N (1998) Breeding system and genetic variance in the monogamous, semi-social shrew, *Crocidura russula*. *Evolution*, **52**, 1230–1235.

Clarke AL, Saether B-E, Roskaft E (1997) Sex biases in avian dispersal: a reappraisal. *Oikos*, **79**, 429–438.

Clutton-Brock TH (1989) Female transfer and inbreeding avoidance in social mammals. *Nature*, **337**, 70–72.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) Comparison of methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics*, **153**, 1989–2000.

Dobson FS (1982) Competition for mates and predominant juvenile male dispersal in mammals. *Animal Behaviour*, **30**, 1183–1192.

Doughty P, Sinervo B, Burghardt GM (1994) Sex-biased dispersal in a polygynous lizard, *Uta stansburiana*. *Animal Behaviour*, **47**, 227–229.

Favre L, Balloux F, Goudet J, Perrin N (1997) Female-biased dispersal in the monogamous mammal *Crocidura russula*: evidence from field data and microsatellite patterns. *Proceeding of the Royal Society of London B*, **264**, 127–132.

Gibbs HL, Dawson RJG, Hobson KA (2000) Limited differentiation in microsatellite DNA variation among northern populations of the yellow warbler: evidence for male-biased gene flow? *Molecular Ecology*, **9**, 2137–2147.

Goldstein DB, Pollock DD (1997) Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of Heredity*, **88**, 335–342.

Gompper ME, Gittleman JL, Wayne RK (1998) Dispersal, philopatry, and genetic relatedness in a social carnivore: comparing males and females. *Molecular Ecology*, **7**, 157–163.

Goudet J, Raymond M, DeMeeus T, Rousset F (1996) Testing differentiation in diploid populations. *Genetics*, **144**, 1933–1940.

Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, **28**, 1140–1162.

Hamilton WD (1971) Selection of selfish and altruistic behavior in some extreme models. In: *Man and Beast: Comparative Social Behavior* (eds Eisenberg JF and Dillon WS), pp. 57–91. Smithsonian Institute Press, Washigton, DC.

Hartl DL, Clark GC (1997) *Principles of Population Genetics*. 3rd edn. Sinauer Associates, Sunderland, MA.

Hölldobler B, Wilson EO (1990) *The Ants*. Springer-Verlag, Berlin.

Ishibashi Y, Saitoh T, Abe S, Yoshida MC (1997) Sex-related spatial kin structure in a spring population of grey-sided voles *Clethrionomys rufocanus* as revealed by mitochondrial and microsatellite DNA analyses. *Molecular Ecology*, **6**, 63–71.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology and Evolution*, **11**, 424–428.

Ji W, Sarre SD, Aitken N, Hankin RKS, Clout MN (2001) Sex-biased dispersal and density-independent mating system in the Australian brushtail possum, as revealed by minisatellite DNA profiling. *Molecular Ecology*, **10**, 1527–1537.

Johnson CN (1986) Sex-biased philopatry and dispersal in mammals. *Oecologia*, **69**, 626–627.

Kalaydjieva L, Calafell F, Jobling MA *et al.* (2001) Patterns of inter- and intra-group genetic diversity in the Vlax Roma as revealed by Y chromosome and mitochondrial DNA lineages. *European Journal of Human Genetics*, **9**, 97–104.

Kawata M (1985) Sex differences in the spatial distribution of genotypes in the red-backed vole, *Clethrionomys rufocanus bedfordiae*. *Journal of Mammalogy*, **66**, 384–387.

Kittles RA, Bergen AW, Urbanek M *et al.* (1999) Autosomal, mitochondrial, and Y chromosome DNA variation in Finland: Evidence for a male-specific bottleneck. *American Journal of Physical Anthropology*, **108**, 381–399.

Knight ME, Van Oppen MJH, Smith HL, Rico C, Hewitt GM, Turner GF (1999) Evidence for male-biased dispersal in lake Malawi cichlids from microsatellites. *Molecular Ecology*, **8**, 1521–1527.

Koenig WD, Van Vuren D, Hooge PN (1996) Detectability, philopatry and the distribution of dispersal distances in vertebrates. *Trends in Ecology and Evolution*, **11**, 514–517.

Manly BJF (1997) *Bootstrap, Randomization and Monte Carlo Methods in Biology*, 2nd edn. Chapman & Hall, London.

McCracken GF (1984) Social dispersion and genetic variation in two species of emballonurid bats. *Zeitschrift Fur Tierpsychologie*, **66**, 55–69.

Melnick DJ, Hoelzer GA (1992) Differences in male and female macaque dispersal lead to contrasting distributions of nuclear and mitochondrial DNA variation. *International Journal of Primatology*, **13**, 1–15.

Mossman CA, Waser PM (1999) Genetic detection of sex-biased dispersal. *Molecular Ecology*, **8**, 1063–1067.

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology*, **4**, 347–354.

Paetkau D, Shields GF, Strobeck C (1998) Gene flow between insular, coastal and interior populations of brown bears in Alaska. *Molecular Ecology*, **7**, 1283–1292.

Part T (1995) The importance of local familiarity and search costs for age- and sex-biased philopatry in the collared flycatcher. *Animal Behaviour*, **49**, 1029–1038.

Perrin N, Goudet J (2001) Inbreeding, kinship and the evolution of natal dispersal. In: *Dispersal* (ed. Clobert J, Danchin E, Dhondt A, Nichols JD), pp. 123–142. Oxford University Press, Oxford.

Perrin N, Mazalov V (1999) Dispersal and Inbreeding Avoidance. *American Naturalist*, **154**, 282–292.

Perrin N, Mazalov V (2000) Local competition, inbreeding, and the evolution of sex-biased dispersal. *American Naturalist*, **155**, 116–127.

Petit E, Balloux F, Goudet J (2001) Sex biased dispersal in a migratory bat: characterization using sex-specific demographic parameters. *Evolution*, **55**, 635–640.

Pritchard JK, Stephens M, Donnely P (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, **155**, 945–959.

Pusey AE (1987) Sex-biased dispersal and inbreeding avoidance in birds and mammals. *Trends in Ecology and Evolution*, **2**, 295–299.

Queller DC, Goodnight KF (1989) Estimating relatedness using genetic markers. *Evolution*, **43**, 258–275.

Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences USA*, **94**, 9197–9201.

Rassmann K, Tautz D, Trillmich F, Gliddon C (1997) The microevolution of the Galapagos marine iguana *Amblyrhynchos cristatus* assessed by nuclear and mitochondrial genetic analyses. *Molecular Ecology*, **6**, 437–452.

Robertson A, Hill WG (1984) Deviations from Hardy–Weinberg proportions: sampling variances and use in estimation of inbreeding coefficients. *Genetics*, **107**, 703–718.

Ross K, Keller L (1995) Joint influence of gene flow and selection on a reproductively important genetic polymorphism in the fire ant *Solenopsis invicta*. *American Naturalist*, **146**, 325–348.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393–430.

Stow AJ, Sunnucks P, Briscoe DA, Gardner MG (2001) The impact of habitat fragmentation on dispersal of Cunningham's skink (*Egernia cunninghami*): evidence from allelic and genotypic analyses of microsatellites. *Molecular Ecology*, **10**, 867–878.

Surridge AK, Ibrahim KM, Bell DJ *et al.* (1999) Fine scale genetic structuring in a natural population of European wild rabbits (*Oryctolagus cuniculus*). *Molecular Ecology*, **8**, 299–307.

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

Wolff JO (1994) More on juvenile dispersal in mammals. *Oikos*, **71**, 349–352.

Wolff JO, Plissner JH (1998) Sex biases in avian natal dispersal: an extension of the mammalian model. *Oikos*, **83**, 327–330.

Zera AJ, Denno RF (1997) Physiology and ecology of dispersal polymorphism in insects. *Annual Review of Entomology*, **42**, 207–231.