Molecular Ecology (2007) **16**, 737–751 doi: 10.1111/j.1365-294X.2006.03184.x

Compatible genetic and ecological estimates of dispersal rates in insect (*Coenagrion mercuriale*: Odonata: Zygoptera) populations: analysis of 'neighbourhood size' using a more precise estimator

PHILLIP C. WATTS,* FRANÇOIS ROUSSET,† ILIK J. SACCHERI,* RAPHAËL LEBLOIS,‡
STEPHEN J. KEMP* and DAVID J. THOMPSON*
**School of Biological Sciences, University of Liverpool, Liverpool L69 7ZB, UK, †Institut des Sciences de l'Evolution, Université Montpellier 2, France, ‡Unité Éco-Anthropologie et Ethnobiologie, Musée de l'Homme, Paris, France*

Abstract

Genetic and demographic estimates of dispersal are often thought to be inconsistent. In this study, we use the damselfly *Coenagrion mercuriale* (Odonata: Zygoptera) as a model to evaluate directly the relationship between estimates of dispersal rate measured during capture–mark–recapture fieldwork with those made from the spatial pattern of genetic markers in linear and two-dimensional habitats. We estimate the 'neighbourhood size' (*Nb*) — the product of the mean axial dispersal rate between parent and offspring and the population density — by a previously described technique, here called the regression method. Because *C. mercuriale* is less philopatric than species investigated previously by the regression method we evaluate a refined estimator that may be more applicable for relatively mobile species. Results from simulations and empirical data sets reveal that the new estimator performs better under most situations, except when dispersal is very localized relative to population density. Analysis of the *C. mercuriale* data extends previous results which demonstrated that demographic and genetic estimates of *Nb* by the regression method are equivalent to within a factor of two at local scales where genetic estimates are less affected by habitat heterogeneity, stochastic processes and/or differential selective regimes. The corollary is that with a little insight into a species' ecology the pattern of spatial genetic structure provides quantitative information on dispersal rates and/or population densities that has real value for conservation management.

Keywords: capture–mark–recapture, conservation, dispersal, isolation by distance, spatial genetic structure

Received 26 August 2006; revision accepted 28 September 2006

Introduction

Given the conservation, ecological and evolutionary significance of dispersal (Clobert *et al*. 2001), much effort has been directed towards quantifying the migration rates of various species. For most organisms, measuring dispersal by direct observation is highly problematic. Consequently, dispersal capabilities are typically assessed using genetic (indirect) techniques. Of some concern, however, are discrepancies between estimates of migration rates provided by direct ecological observations and indirect methods (reviewed by Slatkin 1985; Koenig *et al*. 1996), leading to the opinion that genetic techniques are generally uninformative about levels of contemporary gene flow (Koenig *et al*. 1996; Bossart & Prowell 1998). Certainly where 'numbers of effective migrants' ($N_e m$) are estimated from values of F_{ST}, this discrepancy is largely due to violation of the underlying model assumptions, particularly that of 'island model' (Wright 1931) gene flow whereby all populations exchange equal numbers of migrants (Whitlock & McCauley 1999). One difficulty with the continued axiom of 'ecological-genetic incompatibility' is that only a handful of recent studies have made direct comparisons

Correspondence: Phill Watts, Marine and Freshwater Biology Research Group, The Biosciences Building, School of Biological Sciences, Liverpool University, Crown Street, Liverpool L69 7ZB, UK. Fax: +44 (0)151 795 4408; E-mail: p.c.watts@liv.ac.uk

between direct and indirect estimates of migration rates using newer statistical approaches, and these have produced contrasting results. For example, Spong & Creel's (2001) genetic estimates of dispersal distances were consistent with nearly 6 years of direct observations and Berry *et al.* (2004) obtained reasonably compatible demographic-genetic estimates of interpatch dispersal. By contrast, other studies simply highlight the problems associated with establishing a relationship between direct observations of migration and their genetic counterparts, often because populations have poorly defined boundaries or lack substantial genetic differences (e.g. Adams & Hutchings 2003; Vandewoestijne & Baguette 2004; Wilson *et al.* 2004).

Since the dispersal capabilities of most species are substantially less than their geographical ranges, it is intuitive that neighbours are genetically more alike than distantly separated individuals. This contrasts with the island model and is taken into account by the isolation-by-distance (IBD) models of Wright (1943, 1946) and Malécot (1948). The spatial scale over which IBD develops is proportional to the scale of gene flow, thus suggesting a possible framework to estimate a dispersal rate. Indeed, according to IBD models, one can estimate a 'neighbourhood size' (*Nb*) from the product of the dispersal rate and population density, or, more formally, the mean axial dispersal rate (per generation) between parent and offspring (σ^2) (the mean square parent–offspring distance) and the effective population density (D_e), which is actually a rate of coalescence per unit time and per surface unit (Rousset 1997). In two-dimensions, one generally considers the *Nb* to be equal to $4\pi D_e\sigma^2$. A convenient approximation is that a simple measure of genetic differentiation among individuals (or populations) distributed in two-dimensional space is linearly related to the logarithm of distance (d), i.e. $\approx \ln(d)/(4\pi D_e\sigma^2) + constant$. In one-dimensional (linear) habitats, $Nb = 4D_e\sigma^2$ such that genetic differentiation $\approx d/(4D_e\sigma^2) + constant$ (Rousset 1997). This method of estimation has been discussed in detail elsewhere (Leblois *et al.* 2003, 2004; Vekemans & Hardy 2004). In essence, a simple linear regression of the level of genetic divergence on spatial separation captures information about the combined effect of gene flow and population density, and some independent knowledge of either parameter allows computation of the other. The benefit of this approach is that it may be applied to continuous and discrete populations (Rousset 1997, 2000).

Several studies demonstrate that the regression method (and closely related ones) yields a good correspondence between indirect and direct estimates of *Nb* (Rousset 1997, 2000; Sumner *et al.* 2001; Fenster *et al.* 2003; Winters & Waser 2003; Broquet *et al.* 2006). Despite its analytical simplicity, however, the occurrence of just these few studies implies that this analytical framework is underexploited for the purpose of estimating a dispersal rate. Possible reasons for this include: (i) the model assumption of spatial homogeneity is too restrictive; (ii) failure to detect significant IBD (e.g. Leblois *et al.* 2000); or (iii) that there are still too few simultaneous genetic–demographic assessments of dispersal rates across a variety of taxa using the regression method for it to gain broad acceptance. With respect to the latter it is notable that thorough genetic–demographic comparisons are lacking for insect species (and invertebrates in general) where making accurate estimates of dispersal parameters and population densities can be problematic (Rousset 2004).

Odonates (dragonflies and damselflies) are relatively large and active insects and thus amenable to field-based studies. *Coenagrion mercuriale* (Charpentier) (Odonata: Zygoptera) has emerged as a particularly good model to examine the relationship between direct observations of dispersal and the concomitant pattern of spatial genetic structure in relatively high-density, continuously distributed populations. For example, its UK distribution is well-known, with large populations on Beaulieu Heath, the Itchen Valley (both in southern England) and on the Preseli Hills (southwest Wales) and several smaller colonies elsewhere (Fig. 1). From capture–mark–recapture (CMR) fieldwork, it is evident that most adults move less than 100 m during their

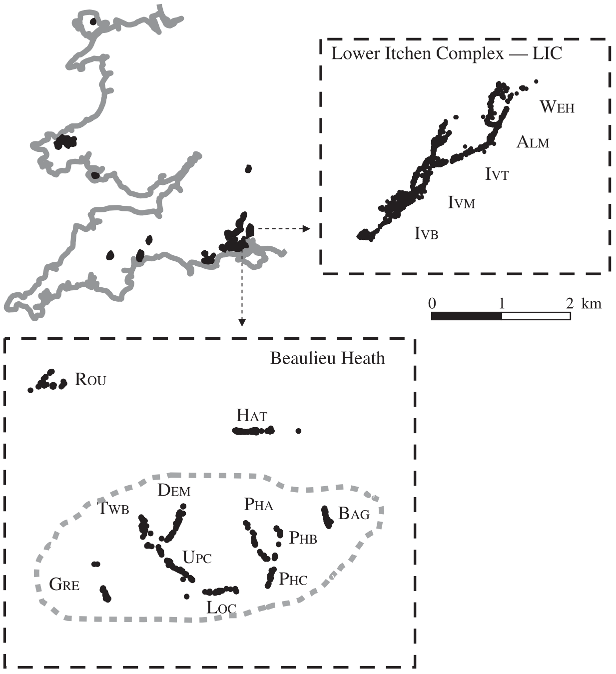

Fig. 1 Approximate distribution of *Coenagrion mercuriale* throughout the UK and the location of the two study sites — the Lower Itchen Complex (LIC) and Beaulieu Heath. Expanded sections display a more precise distribution of *C. mercuriale* throughout each study site. Grey dashed line at Beaulieu Heath indicates the Beaulieu Heath 'continuous' area (see Materials and methods for further details).

lifetimes, although infrequent dispersal over 1 km has been recorded (Hunger & Röske 2001; Purse *et al*. 2003; Watts *et al*. 2004a). Accordingly, in the absence of landscape features that restrict movement, IBD genetic structure develops within large (a few kilometres) habitat patches (Watts *et al*. 2004a, 2006).

There are some potential problems in applying the regression method to the *C. mercuriale* data. First, high mutation has a stronger effect on the accuracy of the method in linear habitats than in two-dimensional ones (Rousset 1997) and this may affect an analysis of riparian systems. Second, the performance of the regression method has been evaluated previously in cases of restricted dispersal that were appropriate for organisms whose dispersal capabilities cover only a few territories (e.g. references in Leblois *et al*. 2004). The $\hat{a}$ estimator of genetic divergence (Rousset's 2000) used in these studies was found by Vekemans & Hardy (2004) to suffer from higher sampling variance than certain measures of 'kinship'; it thus may have a low efficiency for detecting IBD and be poorly suited to analyse dispersal rates in less philopatric species. Accordingly, it appears prudent to evaluate the performance of $\hat{a}$ under higher dispersal rates and consider an alternative test statistic. Statistics that give a higher weight to rare alleles are often better at uncovering IBD but can suffer from bias and thus do not obviously lead to good estimators of *Nb*, as found for Ritland's (1996) estimator by Vekemans and Hardy. However, following a meta-analysis of several data sets, it has been proposed that the statistic of Loiselle *et al*. (1995), which does not give higher weight to rare alleles (explicitly, at least), outperforms $\hat{a}$. An estimator of *Nb* that would accordingly not suffer asymptotic bias has been derived from this statistic (Hardy & Vekemans 1999; Vekemans & Hardy 2004) but its performance in estimation is unknown.

In this study, we (i) use *C. mercuriale* as a model to present a 'worked-example' of the quantitative relationship between ecological observations of dispersal rate and spatial genetic structure in one- and two-dimensional habitats, (ii) reconsider the statistic of Loiselle *et al*. (1995) and derive from it a genetic divergence measure $\hat{e}$ that is analogous to the $\hat{a}$ of Rousset (2000), and (iii) evaluate the performance of this new estimator relative to $\hat{a}$ with regard to more mobile taxa.

We find that the new test statistic $\hat{e}$ suffers from asymptotic bias (which may outweigh its lower variance under certain circumstances), but it nevertheless performs better than $\hat{a}$ for high values of the dispersal rate σ and the upper bound of 95% confidence intervals generated by the ABC bootstrap method are significantly improved over those obtained by the same method but using $\hat{a}$. This allows for accurate analysis of dispersal by *C. mercuriale*, with close agreement between estimates of dispersal rate derived using genetic techniques and fieldwork.

Materials and methods

Description of study sites and field work

Combined genetic–demographic estimates of *Coenagrion mercuriale* dispersal rates were made at Beaulieu Heath (50°47.8′N, 01°29.9′W) and the Itchen Valley (50°57.0′N, 01°20.4′W) (Fig. 1). Beaulieu Heath, isolated from other *C. mercuriale* colonies by more than 4 km of heathland, is a two-dimensional (4.6 × 3.7 km) matrix of seven central (but connected) and four peripheral (and isolated) patches (Fig. 1). At the Itchen Valley, this species inhabits several sites along a 10-km stretch of the River Itchen (Watts *et al*. 2004a). An analysis of *Nb* was possible only at one area — the Lower Itchen Complex (LIC) — because fieldwork was not permitted elsewhere. Nevertheless, the LIC is the largest, continuous area of *C. mercuriale* habitat in this area and is separate from other sites in its vicinity (Watts *et al*. 2004a). The LIC is 2.8 km long and although it attains a maximum width of 629 m is mostly only a few tens of metres wide, so a one-dimensional model of spatial genetic structure is appropriate. For convenience during sampling, the LIC was divided into five areas (Fig. 1) but these do not represent discrete populations.

Adult *C. mercuriale* emerge from May until the end of July, with the peak flight season during June (Purse & Thompson 2003). We undertook CMR for 5 weeks from 12 June 2001 (LIC) and between 11 June and 14 July 2002 (Beaulieu Heath). Adults were searched for every day (09:30–16:00) except during poor weather when they are not active (Banks & Thompson 1985). All unmarked, mature damselflies were caught and marked using the methods described by Watts *et al*. (2004a). When marked animals were observed, their numbers were read using close-focusing binoculars or they were recaptured if there was any doubt as to their number. The position of every encounter was recorded using a differential global positioning system.

Estimation of demographic parameters

Adult lifespan. After emergence, immature adults spend up to 8 days maturing (Purse & Thompson 2003) and are less likely to be observed. Immature individuals were not manipulated in any case because they are susceptible to damage. Adult lifespan was estimated from the mean duration between all first and last encounters and thus represents the mature period. Note that these estimates are provided just for 'ecological context' as the calculations of *Nb* do not depend upon the time interval used (days in this study) but only require that the dispersal rates and rates of coalescence derived from the population densities are measured on the same timescale (see Rousset 1999 for further details).

Adult density. Male *C. mercuriale* are encountered more often than females (Watts *et al.* 2004a); however, genetic and demographic data indicate a 1:1 sex ratio (Corbet 1999; Purse & Thompson 2003; Rouquette & Thompson 2006). The bias towards encountering males reflects differential behaviour, with females only visiting breeding sites when ready to mate while the males, by contrast, are almost always active. Consequently, our CMR data will underestimate female abundance. To overcome this, we used male data to estimate daily population sizes, calculated using a full Jolly–Seber model (Jolly 1965; Seber 1973), which were then doubled to account for the more cryptic females. The numbers of damselflies present outside our sampling (and for days when no CMR was undertaken) were estimated from a logistic growth (or decline) trend based on the available increasing (or declining) daily population estimates and zero adults on the first (or last) date that *C. mercuriale* were sighted in England (6 May and 25 September, D.K. Jenkins personal communication). The censuses at Beaulieu Heath on 23 June, 4 July and 6 July were substantially lower than expected given the overall variation in population size during the breeding season (Fig. 2b), probably because poor weather reduced capture efficiency. Therefore, these censuses were replaced with expected population estimates calculated from the trend in logistic growth or decline. Dividing the sum of all daily censuses by the average lifespan (days) provided a total population estimate (*N*) that was converted to a density (*D*) using the site length or area (m or m^2) calculated using ARCGIS version 8.3 (ESRI Inc., Redlands, CA).

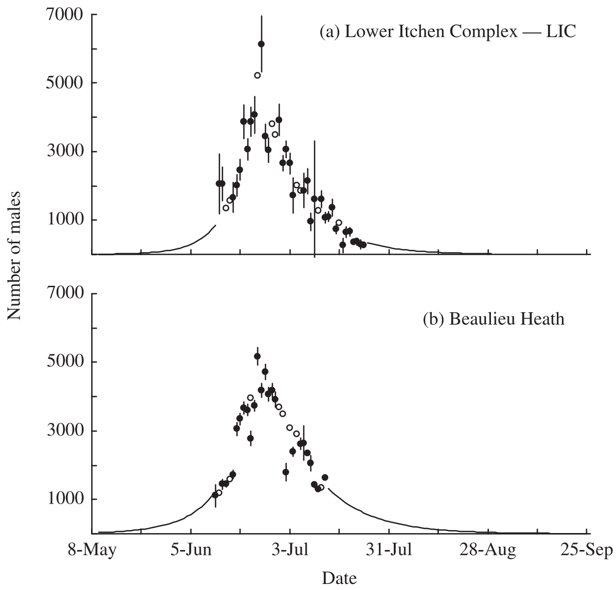

Fig. 2 Variation (± SE) in the estimated daily number of adult male *Coenagrion mercuriale* present at (a) the Lower Itchen Complex (LIC) and (b) Beaulieu Heath. Open circles and solid lines represent data calculated from a logistic growth function that was based on CMR data (solid circles).

Effective population densities (D_e) were calculated as $(N_e/N) \times D$, where N_e is the effective population size and N the adult census. Variance in reproductive success (VRS) among *C. mercuriale* will reduce N_e below N (Frankham *et al.* 2002). To estimate N_e, we employed a Monte Carlo method to determine the VRS of each sex, assuming that reproductive success is directly proportional to lifetime mating success (LMS). Briefly, using Purse & Thompson's (2005) data on the LMS of 77 female and 116 male *C. mercuriale*, the range 0–1 was allocated to individuals in proportion to their calculated LMS. A random number between 0 and 1 was drawn, and the individual whose range contained the random number was then allocated an offspring; this was repeated 193 times (giving a mean number of offspring per individual of 2) and the variance in offspring production then calculated. The procedure was repeated 10 000 times and the average variance calculated. An approximate N_e for each site was calculated using $N_e \approx 8N/(V_{kf} + V_{km} + 4)$, where V_{kf} and V_{km} are the VRS for females and males, respectively (Falconer & Mackay 1996).

Dispersal rate. The 2-year aquatic larval period of *C. mercuriale* (during which they pass through several instars) makes it difficult to estimate a genetic dispersal rate (the distance between a gene and its parent in the preceding generation) using CMR. However, for reasons mentioned in the Discussion, larvae are unlikely to disperse far. Thus, a potential dispersal rate was calculated from the distances moved by adults (our CMR encompassed the entire mature period of most damselflies). The daily average-squared dispersal rate (σ^2) is equal to $0.5 * [\Sigma(dX)^2/\Sigma\, dT]$, where dX is the distance moved by an individual from its release point dT days ago (see Sumner *et al.* 2001 for details), which was then scaled by the estimated average mature adult lifespan.

Genetic analysis

Genotyping. DNA extraction and genotyping methods are described by Watts *et al.* (2004a,b,c). Briefly, genomic DNA was extracted from a tibia for 18–52 damselflies per sample area and every individual was genotyped at 14 unlinked microsatellite loci (see Appendix). Loss of a leg has no measurable effect upon fitness in damselflies (Fincke & Hadrys 2001) and we observed no significant effect of sampling upon recapture rate (D.J. Thompson unpublished).

Data analysis. Every sample (Fig. 1) and the entire data set for each site was tested for departure from expected Hardy–Weinberg equilibrium conditions (HWE) using the permutation test (5000 permutations) in FSTAT version 2.9.3 (Goudet 1995). IBD genetic structure was examined by regression of genetic differentiation among pairs of individuals (or populations) against the distances (or ln-distances in two dimensions) separating them. The measures of genetic differentiation used are estimators of the parameters that obey the theoretical results of Rousset (1997): a multilocus estimator of $F_{ST}/(1-F_{ST})$ between discrete populations, its analogue $\hat{a}$ among pairs of individuals in a continuum (Rousset 2000) or, as explained in the Introduction, another estimator $\hat{e}$ (described below). Individual-based regressions ($\hat{a}$ and $\hat{e}$) were calculated at (i) the LIC, (ii) Beaulieu Heath and also (iii) at Beaulieu Heath but excluding individuals from two sites, ROU and HAT, where neither immigration nor emigration was observed (data not shown). This Beaulieu Heath 'continuous' population (see Fig. 1) thus encompasses a single demographic unit (at least during one breeding season). We also treated the semi-isolated patches at Beaulieu Heath as 'discrete' populations and regressed $F_{ST}/(1-F_{ST})$ against distance to contrast the individual- and population-based estimates of *Nb*. Finally, as the relationship between genetic differentiation and geographical distance is not linear at small scales (Rousset 1997), we repeated all regressions above but excluding pairs of individuals separated by distances less than the demographic estimate of σ (see Table 3). Regressions were made using an upgraded version of GENEPOP 3.4 (Raymond & Rousset 1995) with 95% confidence intervals (CI) about the slopes generated using a nonparametric ABC bootstrap procedure (DiCiccio & Efron 1996; Leblois *et al*. 2003).

To provide some 'conservation' perspective to the quantitative information on dispersal rates that may be derived from genetic-based estimates of *Nb* in combination with demographic data (i.e. an estimate of population density *or* dispersal rate), the mean axial parent–offspring distance (σ) was calculated from the genetic estimate of $4D_e\sigma^2$ or $4\pi D_e\sigma^2$ using the estimate of D_e generated using CMR data; correspondingly, population densities were obtained from the same product using the value of σ^2 provided by the demographic study.

New estimator of the slope, ê

Here, we derive a new estimator from the statistic of Loiselle *et al*. (1995) that is equivalent for testing purposes (and very close for estimation) to the one previously derived by Vekemans & Hardy (2004), but our derivation makes it easier to predict some of its properties. For a pair of individuals i and j, the statistic of Loiselle *et al*. (1995) can be written

$$F = \frac{\sum_k (Y_{ki} - d_k)(Y_{kj} - d_k)}{\sum_k d_k(1-d_k)} - \frac{1}{2n-2}$$

where Y_{ki} is the kth allele frequency in individual i, d_k is the kth allele frequency in the total sample, n is the sample size, and the sums are over all alleles in the sample (Hardy 2003). If we ignore terms that are constant with respect to i and j and thus do not affect the estimation of the slope, then the statistic of Loiselle *et al*. (1995) can be written

$$F = \frac{\sum_k Y_{ki}Y_{kj} - d_k(Y_{ki} + Y_{kj})}{\sum_k d_k(1-d_k)} + \text{constant}$$

$$= \frac{\hat{Q}_{ij} - (\hat{Q}_{i.} + \hat{Q}_{j.})}{\sum_k d_k(1-d_k)} + \text{constant}$$

Table 1 Summary statistics for regression analyses of estimates of genetic differentiation against spatial separation for pairs of individuals ($\hat{a}$ or $\hat{e}$) or populations ($F_{ST}/(1-F_{ST})$) of the damselfly *Coenagrion mercuriale* from the Lower Itchen Complex (LIC) and Beaulieu Heath, UK: all, regression analysis that includes all pairs of individuals; truncated, regression analysis that excludes pairs of individuals within the direct estimate of σ (see Table 2)

Site	Comparison	Estimator	Intercept	Slope	P
Lower Itchen Complex (LIC)	All	$\hat{a}$	-5.54×10^{-2}	4.12×10^{-6}	0.081
		$\hat{e}$	-5.52×10^{-3}	5.58×10^{-6}	< 0.001
	Truncated	$\hat{a}$	-5.57×10^{-2}	3.90×10^{-6}	0.081
		$\hat{e}$	-3.89×10^{-3}	4.49×10^{-6}	< 0.001
Beaulieu Heath	All	$F_{ST}/1-F_{ST}$	-1.10×10^{-2}	2.54×10^{-3}	0.047
		$\hat{a}$	-1.57×10^{-2}	2.59×10^{-3}	0.085
		$\hat{e}$	-1.56×10^{-2}	2.19×10^{-3}	< 0.001
	Truncated	$\hat{a}$	-2.52×10^{-2}	3.86×10^{-3}	0.085
		$\hat{e}$	-1.27×10^{-2}	1.79×10^{-3}	< 0.001
Beaulieu Heath 'Continuous' (i.e. excluding ROU & HAT)	All	$\hat{a}$	-1.53×10^{-2}	2.38×10^{-3}	0.045
		$\hat{e}$	-8.91×10^{-3}	1.33×10^{-3}	< 0.001
	Truncated	$\hat{a}$	-3.58×10^{-2}	5.24×10^{-3}	0.046
		$\hat{e}$	-1.12×10^{-2}	1.65×10^{-3}	< 0.001

P, probability of obtaining a greater correlation than that observed under the null hypothesis (one tailed).

where $\hat{Q}_{ij}$ is the observed identity between individuals i and j, and $\hat{Q}_{i.}$ (resp. $\hat{Q}_{j.}$) is the observed average identity between individual i (resp. j) and all individuals in the sample. To turn it into an estimator of the slope, Vekemans & Hardy (2004) divide it by an approximate estimator of $(1 - Q_w)/\sum_k d_k(1 - d_k)$ where $1 - Q_w$ is the expected frequency of heterozygotes in the sample. We can directly compute

$$\hat{e} \equiv \frac{\hat{Q}_{ij} - (\hat{Q}_{i.} - \hat{Q}_{j.})}{1 - \hat{Q}_w}$$

where $1 - \hat{Q}_w$ is the observed frequency of homozygotes, and the slope from the regression of $\hat{e}$ with (logarithm of) geographical distance should have essentially the same properties as the slope estimator of Hardy & Vekemans. Accordingly, for six simulated data sets, the $\hat{e}$ slopes were found to differ only by $\approx 1/(2n)$ from the slopes derived from the statistic of Loiselle *et al.* (1995) using SPAGEDI (Hardy & Vekemans 2002), where n is the sample size.

Here the main effect of the term $\hat{Q}_{i.} + \hat{Q}_{j.}$ is to decrease the genetic similarity measure $\hat{e}$ when the pair of individuals harbours alleles that tend to be common in the total population, thereby giving more weight to rare alleles in the measurement of genetic similarity. This is expected to reduce the variance of such estimators, though possibly introduce some bias. By comparison, the estimator of Rousset (2000) infers the slope from the variation of $Q_{ii}/(1 - Q_w)$ with distance. This differs from $\hat{e}$ only by the $\hat{Q}_{i.} + \hat{Q}_{i}$ term in the numerator. As $1/Nb$ is the slope of $Q_{ij}/(1 - Q_w), Q_{ij}/(1 - Q_w)$ provides an asymptotically unbiased estimator of Nb. By contrast, the $\hat{Q}_{i.} + \hat{Q}_{j.}$ term in $\hat{e}$ should introduce a bias. For example, individuals at opposite edges of the sampled area are on average more distant in space from random individuals in the sample than are individuals taken in the centre of the sampled range, and thus pairs i, j involving the most distant individuals should tend to have lower $\hat{Q}_{i.} + \hat{Q}_{j.}$ and thus appear more similar than implied by the unbiased estimator. In other words, the divergence between the most distant individuals should be underestimated, thereby lowering the slope estimate, and the more so the stronger the spatial patterns. However, for reasonably sized samples such a bias may be compensated for by a lower variance and simulations will be used to compare the overall performance of the two estimators for Nb estimation. They will confirm that $\hat{e}$-based estimates of the slope tend to be downward biased (i.e. Nb overestimated) but that they have lower mean square error (MSE) unless dispersal is very localized (methods of calculating the relative bias and the MSE are detailed in Leblois *et al.* 2003, 2004).

Data simulations

Our initial aim was to check that the $\hat{e}$-based estimator was better (in some respects) in the context of the LIC data, so the first simulations closely matched this context. Then dispersal was varied to check our understanding of this estimator performance as predicted above. In particular, lower dispersal values were considered to also allow comparison with previous simulation studies of the $\hat{a}$ estimator. Similar considerations guided the two-dimensional simulations, except that we did not try to match closely the more complex distribution of samples in that case. Finally, the effect of immigration from an external source was simulated to address concerns about the behaviour of the estimators in that case. A detailed description of the sample-generating simulation program can be found in Leblois *et al.* (2003, 2004). The main differences in this study are the range of σ^2 values considered and a test of the effect of additional long-distance immigrants. First, we used a family of dispersal distributions obtained as mixtures of convolutions of stepping-stone steps as a convenient way to model discrete distributions with various forms (Chesson & Lee 2005). As detailed in that study, the Sichel mixture is described by three parameters, ξ, ω and γ. We used the long-tailed variant of this family, which is obtained in the limit case $\omega \to 0$, $\xi \to \infty$ with $\omega\xi \to \kappa$. The two parameters γ and κ then describe a family of distributions which are Gaussian-looking at short distances but have tails proportional to $r^{-2\gamma-1}$ for distance r. The values of γ and $\hat{e}$ were chosen so as to achieve given σ and kurtosis for the unbounded dispersal distribution. Second, immigration from a large distant source brings unrelated genes, analogous to the effect of mutation. Hence, 1% and 0.1% immigration rates of individuals from a distant source were simulated by assuming mutation rates of 1% and 0.1%. In other cases, the mutation rate was set as described below.

A linear array of 3500 demes of four diploid individuals was simulated — each deme thus representing about 1 m of the LIC habitat. Under conditions of (i) relatively high dispersal ($\sigma = 25$ & 130 lattice steps), 240 individuals were sampled at a density of one individual every 11 demes (which matches with the density of sampling in the LIC), while (ii) for more limited dispersal ($\sigma = 5$) 100 individuals were sampled at a rate of one individual from every deme. Previous simulations (Leblois *et al.* 2003) have shown that genetic diversity is a major determinant of the performance of estimation, so the mutation rate (μ) was chosen to achieve levels of diversity similar to those found in the LIC (heterozygote frequency = 0.57). For $\sigma = 130$, we thus chose $\mu = 4 \times 10^{-5}$ and retained this value in further simulations, except in some cases where it resulted in too high diversity (see Table 3).

In two dimensions, a 500 × 500 lattice with absorbing boundaries was simulated, comparable to the sampling density of the simulated linear habitat, with independent dispersal in each dimension. We considered the cases: (i) $\sigma = 5$, with four diploid individuals per deme and 225 individuals sampled on a 15 × 15 grid over a 43 × 43 surface, one individual being sampled every three steps in each dimension; (ii) same but with one diploid individual per lattice node; and (iii) same as (ii) but with $\sigma = 3$ and (4) $\sigma = 2$, with one individual per deme and 100 individuals sampled on a 10 × 10 surface. With higher dispersal, efficient estimation of *Nb* becomes difficult.

Results

Demographic parameters

Adult lifespan. These two CMR studies, the largest undertaken so far for any odonate, involved thousands of marked/recaptured damselflies — 10 259/4158 and 6783/1747 at Beaulieu Heath and the LIC, respectively. Low recapture rates are typical for CMR studies of odonates (Corbet 1999) and likely reflect the short mature adult lifespan. Some individuals were observed over a period of several weeks following marking, but the average duration (± SE) between first and last captures was 5.11 (± 0.10) days at the LIC and 5.93 (± 0.07) days at Beaulieu Heath.

Adult density. At both sites, the number of individuals on any day during the peak flight period was considerable, attaining a maximum of 5000–6000 males (Fig. 2a, b). Overall, the respective population estimates at the LIC and Beaulieu Heath are (approximately) 37 868 and 39 913 damselflies. From the proportion of marked individuals (data not shown), the isolated Rou and Hat sites comprise about 19% of the total population at Beaulieu Heath. Adult densities at LIC, Beaulieu Heath and the Beaulieu Heath 'continuous' populations are estimated at 13.18 individuals m^{-1}, 4.45×10^{-3} individuals m^{-2} and 8.82×10^{-3} individuals m^{-2}. Estimated VRS of 7.4 (female) and 13.5 (male) provide a N_e/N ratio of 0.32 and the scaled D_e's in Table 2.

Dispersal rate. A similar pattern and scale of movement was evident at both sites, with *Coenagrion mercuriale* not dispersing freely throughout either habitat matrix. Just over 75% of adults moved less than 100 m, while 95% of adults were found within 300 m of their initial mark site (cumulative distance moved over all recaptures) (Fig. 3a, b). Mean (cumulative) lifetime distance moved (± SE) was 89.88 (± 3.78) m at the LIC and 87.33 (± 2.20) m on Beaulieu Heath. The respective estimates of daily mean axial dispersal rate (σ^2) at LIC, Beaulieu Heath and the Beaulieu Heath 'continuous' populations of 3214.90 m^2, 2335.16 m^2 and 2768.78 m^2 provide corresponding demographic *Nb* estimates of 277 894 individuals m^{-1}, 249 individuals and 555 individuals (Table 2).

Table 2 Comparison of demographic-(CMR) and genetic-(microsatellite) based methods of estimating neighbourhood size ($Nb = 4D_e\sigma^2$ or $4\pi D_e\sigma^2$ for one- or two-dimensional habitats, respectively), dispersal distance (σ) and effective population density (D_e) in the damselfly *Coenagrion mercuriale* from the Lower Itchen Complex (Itchen Valley) and on Beaulieu Heath (New Forest), UK

1-D (one-dimensional) 2-D (two-dimensional)	Estimator	*Nb* (individuals)	95% CI of *Nb* (individuals)	σ (m)	D_e (ind*m^{-1}) (ind*m^{-2})
Lower Itchen Complex (LIC)					
Direct estimate		277 894		128.11	4.23
Indirect estimate	$\hat{a}$	242 816	66 015–8	119.75	3.70
	$\hat{e}$	179 058	76 949–392 866	102.87	2.73
Indirect estimate[1]	$\hat{a}$	256 498	51 324–8	123.08	3.91
	$\hat{e}$	222 666	88 966–546 245	114.72	3.39
Beaulieu Heath (all sites)					
Direct estimate		249		117.72	1.43×10^{-3}
Indirect estimate	$F_{ST}/(1 - F_{ST})$	393	93–8	147.88	2.26×10^{-3}
	$\hat{a}$	386	143–8	146.55	2.22×10^{-3}
	$\hat{e}$	456	196–1827	159.30	2.62×10^{-3}
Indirect estimate[1]	$\hat{a}$	259	86–8	120.04	1.49×10^{-3}
	$\hat{e}$	557	242–2348	176.06	3.19×10^{-3}
Beaulieu Heath 'continuous' sites (i.e. excluding ROU & HAT)					
Direct estimate		555		125.06	3.00×10^{-3}
Indirect estimate	$\hat{a}$	421	178–8	108.69	2.14×10^{-3}
	$\hat{e}$	753	319–3162	141.33	3.83×10^{-3}
Indirect estimate[1]	$\hat{a}$	191	75–8	73.23	9.71×10^{-4}
	$\hat{e}$	606	177–17 097	126.79	3.08×10^{-3}

[1] Indirect estimate made using only pairs of individuals separated by distances greater than the direct estimate of σ (i.e. truncated regression).

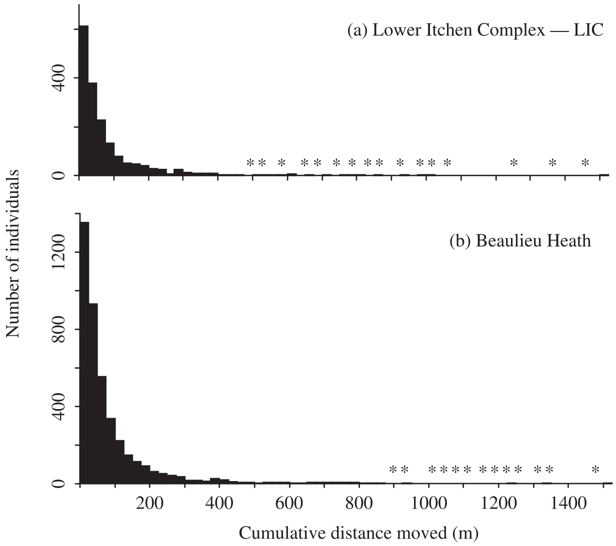

Fig. 3 Frequency of cumulative lifetime movement of adult *Coenagrion mercuriale* in 25-m distance categories for (a) the Lower Itchen Complex (LIC) and (b) Beaulieu Heath, both to the same scale. *n*, number of recaptured individuals; *highlights infrequent ($n = 1$ or 2) movement events.

Pattern of genetic differentiation

We genotyped 240 and 489 individuals at the LIC and Beaulieu Heath, respectively; sample sizes and basic indices of genetic diversity are given in the Appendix. All but five (two at LIC, three at Beaulieu Heath) of the 224 locus-sample combinations met ($P > 0.05$) expected HWE conditions after a sequential Bonferroni correction (Rice 1989) applied for $k = 14$ loci per sample (Appendix). Global tests for HWE (all samples within a site combined) revealed significant ($P < 0.05$, $k = 14$) excesses of homozygotes at LIST4-023 & LIST4-060 at the LIC and at LIST4-066 at Beaulieu Heath. Overall, the signal of departure from random mating within samples and sites is minimal and all loci were retained for the regression analyses.

The considerable variability in the observed relationship between pairwise genetic differentiation and spatial separation (Fig. 4a–c) is a general feature of mutation-drift models. Nonetheless, positive regression slopes were produced for all analyses (Table 1, Fig. 4a–c), with a contrast between gradients based on $\hat{e}$ that were all significantly different ($P < 0.05$) from the null slope and those generated using $\hat{a}$, which were at best 'weakly' significant ($P \approx 0.045$, Beaulieu Heath 'continuous'). Full details of all analyses are provided (Tables 1 and 2) but in the following section we concentrate on the results for the new estimator $\hat{e}$.

Regression of $\hat{e}$ against geographical distance for all pairs of individuals in the LIC gave a slope of 5.58×10^{-6} that is equivalent to *Nb* = 179 058 individuals m^{-1} (Tables 1 and 2). On Beaulieu Heath and Beaulieu Heath 'continuous', gradients of 2.19×10^{-3} and 1.33×10^{-3} provide *Nb* estimates of 456 and 753 individuals (Tables 1 and 2). Truncated regressions generated comparable estimates of *Nb*: LIC = 222 666 individuals m^{-1}, Beaulieu Heath = 557 individuals and Beaulieu Heath 'continuous' = 606 individuals. A population-based regression at Beaulieu Heath provided an *Nb* estimate of 393 individuals, which is comparable to that based on individual genetic differences (Table 1). With one exception (Beaulieu Heath 'continuous'), regressions of $\hat{e}$ on geographical distance provided equivalent (to within a factor of two) estimates of *Nb* compared with those made using demographic (CMR) data, albeit with a slightly smaller *Nb* in the (one-dimensional) LIC and an increased *Nb* at (two-dimensional) Beaulieu Heath. All 95% CIs of the genetic estimates of *Nb* based on $\hat{e}$ enclosed the corresponding demographic estimate of *Nb*, but the ABC bootstrap procedure failed to provide a finite upper CI to slopes based on $\hat{a}$ or F_{ST} (Tables 1 and 2). Because of the close agreement between the genetic and demographic estimates of *Nb*, a calculation of effective dispersal rate (σ) that combines the indirect estimate of *Nb* and the demographic estimate of (effective) density is comparable to that made solely using demographic data. Likewise, values for D_e estimated using demographic estimates of dispersal and genetically derived *Nb* are similar to those based on fieldwork (Table 2).

Performance of new estimator

The main result is that the estimators of *Nb* are generally biased upwards (i.e. slope estimates are biased downwards) (Table 3); an obvious explanation for this is that mutation reduces differentiation. Theoretical predictions of the magnitude of the bias due to mutation may not be useful in practice because they require an estimate of mutation rate and σ, but they serve to understand the simulation results; thus, the performance of estimation of *Nb* corrected for the effect of mutation is also presented in the two cases with largest bias (Table 3). Although some approximations are available to predict the bias (Rousset 1997) they do not accurately take into account high mutation rates and edge effects, hence the expected slope was approximated by simulation of 40 000 independent loci.

From a comparison of these results in a linear habitat, it appears that the $\hat{a}$-based estimator of Rousset (2000) usually has lower bias than the $\hat{e}$-based one, although this discrepancy tends to diminish with increasing dispersal rate. However, the relative root MSEs for both estimators are similar when dispersal is most limited but not for the case of $\sigma = 130$ where the relative MSE is substantially greater for $\hat{a}$ than $\hat{e}$. Generally, in two-dimensional space the relative root MSEs are greater for $\hat{a}$ than $\hat{e}$, particularly for a higher dispersal rate (Table 3). The accuracy of the upper bound of the CI for *Nb* follows the same

Table 3 Relative performance of two estimators of genetic differentiation ($\hat{a}$ and $\hat{e}$) to detect IBD under different rates of dispersal

									Slope (1/*Nb*)		*Nb*			
γ	$\hat{e}$	σ	Kurtosis	μ	*H*	*D*	*n*		Relative bias	Relative ROOT MSE	CI too high	CI too low	Estimate < 0	CI contains 0
One-dimensional space (linear habitat)														
−5.13	206.25	5	1	4×10^{-5}	0.378	4	100 (1)	$\hat{a}$	−0.071	0.333	0.115	0.025	0.000	0.015
								$\hat{e}$	−0.262	0.350	0.260	0.005	0.000	0.000
−2.15	57.52	5	20	4×10^{-5}	0.396	4	100 (1)	$\hat{a}$	−0.064	0.309	0.090	0.015	0.000	0.015
								$\hat{e}$	−0.247	0.343	0.265	0.000	0.000	0.000
	1437.52	25	20	4×10^{-5}	0.608	4	240 (11)	$\hat{a}$	−0.246	0.300	0.375	0.000	0.000	0.000
								$\hat{e}$	−0.352	0.377	0.585	0.000	0.000	0.000
Performance relative to slope from 40 000 loci								$\hat{a}$	0.038	0.239	0.055	0.060		
								$\hat{e}$	−0.108	0.217	0.180	0.000		
	38 870.02	130	20	4×10^{-5}	0.645	4	240 (11)	$\hat{a}$	−0.112	0.922	0.055	0.035	0.200	0.820
								$\hat{e}$	−0.241	0.429	0.205	0.005	0.000	0.125
Performance relative to slope from 40 000 loci								$\hat{a}$	0.022	1.053	0.035	0.035		
								$\hat{e}$	−0.127	0.428	0.135	0.015		
				10^{-3}	0.883	4	240 (11)	$\hat{a}$	−0.218	0.528	0.115	0.020	0.025	0.610
								$\hat{e}$	−0.318	0.368	0.405	0.000	0.000	0.000
Simulated effect of migration				10^{-2}	0.897	4	240 (11)	$\hat{a}$	−0.635	0.790	0.400	0.000	0.250	0.805
								$\hat{e}$	−0.649	0.665	0.930	0.000	0.000	0.145
Two-dimensional space														
−2.15	9.22	2	20	4×10^{-5}	0.810	1	100 (1)	$\hat{a}$	−0.100	0.237	0.095	0.010	0.000	0.005
								$\hat{e}$	−0.271	0.302	0.640	0.000	0.000	0.000
				5×10^{-7}	0.369	1	100 (1)	$\hat{a}$	−0.093	0.543	0.080	0.025	0.025	0.515
								$\hat{e}$	−0.259	0.398	0.275	0.010	0.000	0.075
	20.72	3	20	5×10^{-7}	0.352	1	225 (3)	$\hat{a}$	−0.044	0.701	0.030	0.035	0.090	0.645
								$\hat{e}$	−0.130	0.340	0.160	0.020	0.000	0.060
				10^{-3}	0.873	1	225 (3)	$\hat{a}$	−0.077	0.219	0.075	0.025	0.000	0.010
Simulated effect of migration								$\hat{e}$	−0.208	0.238	0.465	0.000	0.000	0.000
				10^{-2}	0.883	1	225 (3)	$\hat{a}$	−0.343	0.390	0.465	0.000	0.000	0.055
								$\hat{e}$	−0.411	0.419	0.970	0.000	0.000	0.000
	57.52	5	20	5×10^{-7}	0.344	1	225 (3)	$\hat{a}$	−0.020	1.839	0.060	0.060	0.260	0.900
								$\hat{e}$	−0.096	0.640	0.110	0.025	0.025	0.435
				4×10^{-5}	0.895	4	225 (3)	$\hat{a}$	−0.102	1.787	0.035	0.050	0.290	0.900
								$\hat{e}$	−0.123	0.597	0.035	0.025	0.045	0.675
				5×10^{-7}	0.615	4	225 (3)	$\hat{a}$	0.033	3.623	0.050	0.030	0.370	0.935
								$\hat{e}$	−0.044	1.121	0.080	0.030	0.210	0.825

γ & $\hat{e}$, parameters describing dispersal distribution; σ, dispersal rate; μ, mutation rate; *H*, heterozygote frequency; *D*, density; *n*, number of individuals sampled (in parentheses: sampling rate of individuals on lattice, see Methods); MSE, mean square error; CI, 95% confidence interval of slope. Note that ideally, for 95% CI's the frequency of CI too low or too high should both be 0.025.

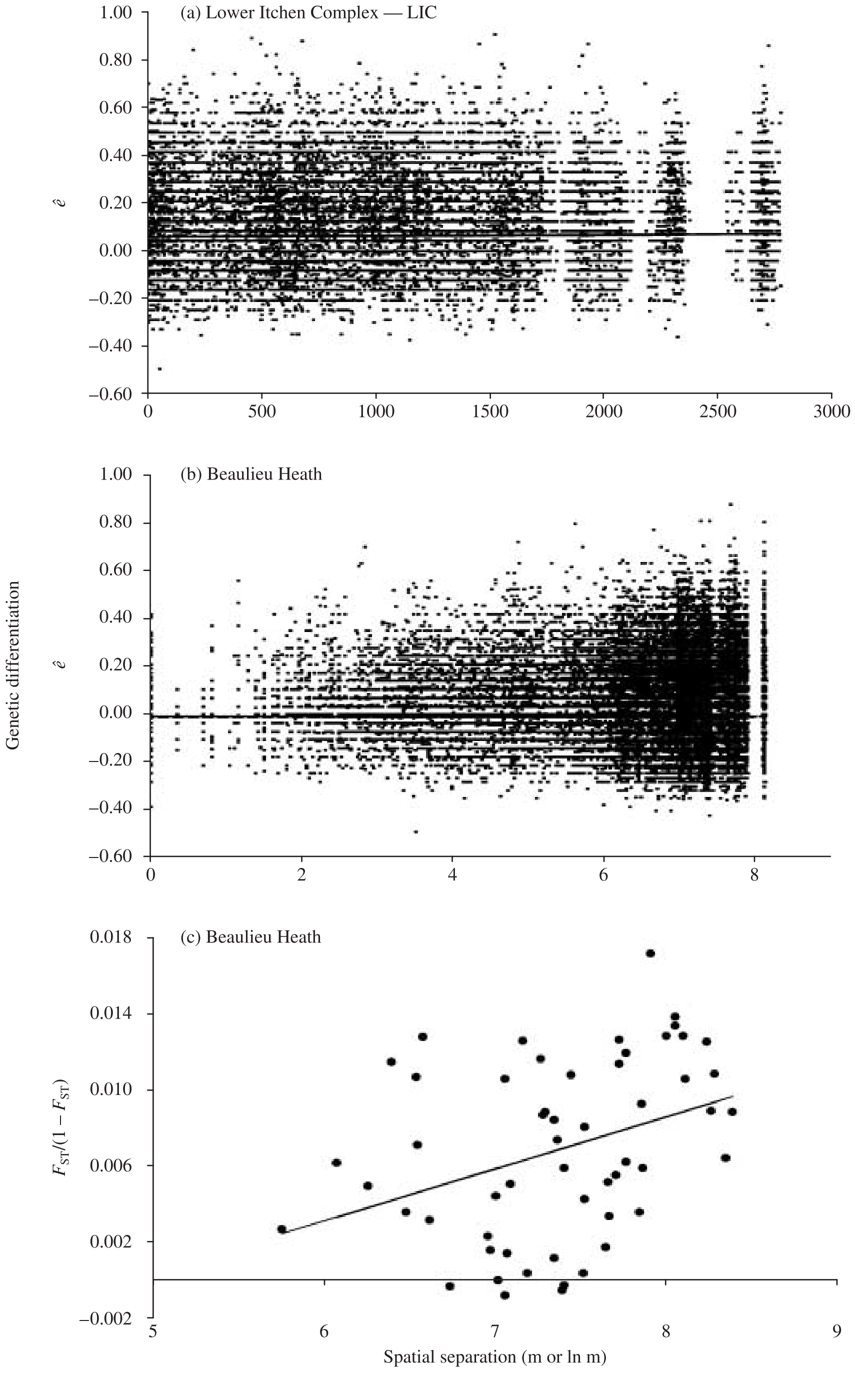

Fig. 4 Linear regression between the geographical distance separating pairs of individual *Coenagrion mercuriale* in (a) the Lower Itchen Complex (LIC) and (b) Beaulieu Heath and the corresponding estimate of pairwise genetic differentiation ($\hat{e}$). Also shown is (c) the relationship between spatial separation and level of pairwise genetic differentiation (defined by $F_{ST}/[1 - F_{ST}]$) among pairs of samples at Beaulieu Heath. Note that distances at the two-dimensional habitat, Beaulieu Heath, are provided as ln-metres.

trend ('CI too low', Table 3) while $\hat{a}$ provides consistently more accurate lower bounds trend ('CI too high').

With immigration from a large distance source, estimation of the local dispersal rate degrades (i.e. increased negative relative bias, Table 3). In the cases presented here, with a 0.01 immigration rate the average slope is only one-third of the value expected from the local component of dispersal in the linear case, and ~60% of this expected value in two dimensions. In contrast to $\hat{a}$, the new estimator $\hat{e}$ still retains a high power (> 0.85) to detect IBD (i.e. null slope is not included in CI) and MSE can actually be improved by such immigration through its effect on gene diversity.

Discussion

The view that spatial genetic structure does not reflect the pattern of contemporary gene flow is commonly put forward, largely because of a putative confounding effect of historical patterns of gene flow. Certainly this is true for many populations separated by large distances where ecological movement is probably irrelevant. In this study, an assessment of two extensive simultaneous demographic–genetic data sets demonstrates that the pattern of spatial genetic structure provides an estimate of *Nb* (Rousset 1997) that is equivalent (within a factor of two or better) to that obtained from ecological observations (Table 2). This level of accuracy is consistent with that observed by previous studies using the regression method (cited in the Introduction) and is expected from simulation studies (Leblois *et al*. 2003). Given that many studies have failed to find this correlation, why are the results from this method compatible?

Comparison between direct and indirect estimates of dispersal rate

Correspondence between direct and indirect estimates of dispersal rates relies upon a minimal impact of a number of possible confounding factors. For example, handling may invoke increased movement away from the site of disturbance. Potentially more problematic, however, is the measurement of total (cumulative) dispersal by mature adults and not that of genes (see Methods). Numerous species travel over long distances but return to distinct areas to reproduce. Moreover, immigrants may experience low reproductive success, for example because of a cost to dispersal or from poor adaptation to local conditions (Marr *et al*. 2002; Hansson *et al*. 2004). Thus, direct observations of dispersal can overestimate the movement of genes. While these factors do not appear to be significant in this study (cf. Table 2), an explicit investigation into the contrast between breeding and foraging movements or the reproductive cost of dispersal has not been undertaken for any odonate.

Generally, field studies are expected to negatively bias estimates of dispersal rates because of movement during unsampled life-history stages (Mallet 1986; Wilson *et al*. 2004), infrequent, long-distance dispersal (Slatkin 1985) and/or spatially restricted sampling (Koenig *et al*. 1996; Hanski 2003; Schneider 2003). There is the prospect for pre-reproductive movement during the 2 years larval stage or as an immature adult. Larval drift is a feature of many freshwater invertebrates (Elliott 2003) but considered unlikely for *Coenagrion mercuriale* larvae that inhabit shallow, slow flowing watercourses and are thigmotactic. Similarly, while immature adult dispersal has been observed in some damselflies (Banks & Thompson 1985; Corbet 1999), this behaviour has not been documented during our fieldwork. Finally, our adult dispersal distributions are unlikely to be appreciably truncated because both sites are surrounded by large areas of inhospitable habitat, and these and other studies have documented that *C. mercuriale* do not disperse more than 2 km even though the study areas (and suitable habitat) extend farther (Hunger & Röske 2001; Purse *et al*. 2003).

The effective population sizes of most natural populations are typically less than the adult censuses, through some combination of demographic fluctuations, VRS or uneven sex ratio (Falconer & Mackay 1996; Frankham *et al*. 2002). Genetic, behavioural and demographic data (Corbet 1999; Purse & Thompson 2003; Rouquette & Thompson 2006) indicate that the latter has little or no effect in lowering the N_e of *C. mercuriale,* and accordingly, we estimated N_e from the expected effect of VRS only. However, over many generations the sizes of *C. mercuriale* populations are likely to vary, like those of other insects (Hanski 2003; Gardarsson *et al*. 2004). With relevant ecological data (which do not exist for *C. mercuriale*) to take this into consideration, the N_e/N ratio would be reduced further, possibly to the extent that the demographic estimates of *Nb* would fall below the lower 95% CI of the indirect estimates of *Nb*. Accepting our genetic estimates of *Nb* to be relatively unbiased implies that a single N_e/N ratio which incorporates demographic fluctuations as well as factors operating every year (VRS, sex ratio or age structure) may not be very informative. With this in mind, it is also relevant that there is a better correspondence at the LIC and Beaulieu Heath 'continuous' sites compared with Beaulieu Heath. This may reflect an effect of scale (i.e. more localized sampling) whereby habitat continuity, higher dispersal rates and/or more routine gene flow increases the rate of approach to genetic equilibrium conditions (Slatkin 1993; Hardy & Vekemans 1999; Rousset 2006), hence the agreement between ecological and genetic dispersal estimates. Overall, the good correspondence between genetic and demographic estimates of *Nb* suggests a minimal impact of the various factors discussed above.

It has often been assumed that a small rate of long-distance immigration, easily missed by demographic studies, would have a disproportionate effect on genetic patterns and would explain discrepancies between genetic and demographic estimates. On the contrary, the regression method is based on genetic patterns which are robust to such immigration, and consequently, good matches between demographics analyses of local dispersal and local genetic differentiation are expected. This correspondence is expected to degrade as the long-distance immigration rate increases, in a predictable way (see Rousset 1997 and p. 42 of Rousset 2004 for predictions in terms of μ, σ, and distance between samples). We presented some simulations for illustration, where the effect on the regression method would be moderate when individuals have less than a 0.1% probability of being long-distance immigrants missed by the demographic study (Table 3).

Performance of estimators

In a linear habitat, the $\hat{a}$-based estimator has lower bias than the $\hat{e}$-based one, and when dispersal is limited, the

MSEs of both estimators are similar. For short-distance sampling, mutation has little effect (i.e. modest bias). Thus, when the effect of mutation is taken into account, we see that (i) the bias of $\hat{a}$-based estimates is largely removed but the lower bound of their CIs tends to be too high, as previously noted by Leblois *et al*. (2003), and (ii) a bias remains for $\hat{e}$-based estimates and their CIs are correspondingly inaccurate. However, at larger scales ($\sigma = 130$) the issue of bias become less important than the MSE, and consequently, the slope estimates of the $\hat{a}$-based estimator are often negative and the CI more frequently includes a null slope than the $\hat{e}$-based estimator, which is consistent with the putative higher testing power of the statistic of Loiselle *et al*. (1995). Overall, we may expect $\hat{e}$ to be superior when the spatial pattern is weak (as a result of large σ relative to distances among individuals) while $\hat{a}$ appears superior given the opposite. In two dimensions, the trends are similar but the effects of mutation are less apparent, partly because it is easier to sample as many individuals within a smaller maximum distance compared with linear space. That the $\hat{e}$ bias is smaller in two dimensions for identical σ-value is expected since the expected pattern of IBD is weaker in two dimensions (see predictions of bias of the estimator in the Methods). Overall, $\hat{e}$ had lower MSE than $\hat{a}$ and appears to be the statistic of choice except where dispersal is restricted ($\sigma = 2$). Accordingly, a rough rule of thumb would be to perform the $\hat{e}$-based analysis in all cases, and if it yields low estimates ($Nb < 10\,000$ in one dimension, < 50 in two dimensions), to perform the $\hat{a}$-based one in order to obtain better point estimates. Since $\hat{a}$ consistently provides more accurate lower bounds for the CI, in most cases one should derive the *Nb* lower bound from the $\hat{a}$-based analysis and the upper bound from the $\hat{e}$-based one.

The regression method based on the $\hat{e}$ statistic shares properties of its previous implementation based on the $\hat{a}$ statistic. In particular, accurate estimates of *Nb* may be obtained with a leptokurtic pattern of dispersal (Fig. 2a-c) which is expected given the weak assumptions made by the demographic model about the distribution of dispersal distances (Rousset 1997, 2000; Leblois *et al*. 2003). However, $\hat{e}$ alleviates limitation of the $\hat{a}$-based estimator in estimating spatial patterns in relatively mobile species ('mobility' refers to the number of territories moved rather than distance per se). Although the simulation study is limited, its results are consistent with the predictions of bias and also with differences in power suggested by the analysis of several data sets (Vekemans & Hardy 2004). For *C. mercuriale*, we find the new estimator $\hat{e}$ to be superior at detecting genetic structure than $\hat{a}$ where high variance likely explains the failure of ABC bootstrap to provide a finite upper CI (Table 3). Likewise, for an invasive cane toad population, Leblois *et al*. (2000) reported a large (90 to infinity) CI for *Nb*. Reanalysis of these data using $\hat{e}$ yields the point estimate 232 with CI 125–1205 (the Mantel test remains nonsignificant, one-tailed $P = 0.124$).

Moreover, the $\hat{e}$-based analysis should be relatively independent from many past demographic events much as the $\hat{a}$-based one (Table 7 in Leblois *et al*. 2004). Both methods assume genetic equilibrium, a condition that is approached more rapidly at local scales (Slatkin 1993; Hardy & Vekemans 1999; Rousset 2006), and it is important to limit sampling to within about 10–50 times σ (Rousset 2000; Vekemans & Hardy 2004) or less if mutation rates are particularly high (Rousset 1997): an appropriate scale of analysis for *C. mercuriale* thus lies between 1.3 km and 6.5 km. In addition, although the relationship between genetic differentiation and distance is not linear at distances less than the demographic estimate of σ (Rousset 1997) we note that this bias (as computed from expected patterns of genetic structure) associated with incorporating all data is weak and accordingly the truncated regression generally differs little from the full data analysis (Table 3, see also Rousset 2000; Sumner *et al*. 2001).

Finally, it is worth noting that the statistical power of the population-based analysis is weaker than one based on genetic differences among individuals, although both methods yielded similar point estimates of *Nb*. Relatively poor statistical power may account for an apparent absence of IBD in species where CMR indicates that dispersal is spatially restricted (Castric & Bernatchez 2004).

Conservation implications

A large effort has been directed to the conservation of *C. mercuriale* populations in the UK. We have shown that *C. mercuriale* do not disperse freely throughout a habitat matrix of several kilometres, thereby emphasizing the importance of maintaining habitat continuity even at local scales to prevent population fragmentation and accumulation of genetic differences and loss of diversity that this species is apparently prone to (Watts *et al*. 2005, 2006). More significant, is that we are able to reconcile demographic estimates of dispersal with the contemporary pattern of spatial genetic structure and provide some guide to the appropriate test statistics for both high and low dispersal species. Moreover, failure to detect IBD implies an inadequate spatial scale of sampling that is informative itself, for example by providing evidence for abundant long-distance dispersal (with respect to the analysis). Thus the regression method has a real conservation value by quantifying an ecologically relevant dispersal rate that can be integrated into management plans.

Summary

The regression-based approach to quantifying dispersal benefits from being analytically straightforward, robust to deviations from the model assumptions and applicable to discrete populations as well as individuals within a continuum. Accordingly, we find that using an appropriate estimator, demographic and genetic estimates of 'neighbourhood size' are equivalent to within a factor of two at local scales. In order to increase the overall precision in estimating a dispersal rate, we presented a new statistic $\hat{e}$ and examined its performance relative to a previously derived test statistic $\hat{a}$. Empirical data and computer simulations reveal trends that are consistent with theoretical expectations; that is, $\hat{e}$ is better test statistic and a better estimator under many situations, but the previously used statistic $\hat{a}$ remains appropriate when dispersal is localized relative to population density.

Acknowledgements

Coenagrion mercuriale is protected under Schedule 5 of the Wildlife & Countryside Act (1981). All work was carried out under licence from English Nature. We thank all the landowners for allowing us onto their land. We are grateful to the NERC (NER/A/S/2000/01322), the Itchen Sustainability Study Group and the Environment Agency for provision of funds. We thank Tim Sykes, Alison Strange and all those involved in the CMR studies for their help. Ian Harvey wrote the program that performed the scaling procedure. We also acknowledge the comments made by several reviewers that helped improve this manuscript.

References

Adams BK, Hutchings JA (2003) Microgeographic population structure of brook charr: a comparison of microsatellite and mark–recapture data. *Journal of Fish Biology*, **62**, 517–533.

Banks MJ, Thompson DJ (1985) Lifetime mating success in the damselfly *Coenagrion puella*. *Animal Behaviour*, **33**, 1175–1183.

Berry O, Tocher MD, Sarre SD (2004) Can assignment tests measure dispersal? *Molecular Ecology*, **13**, 551–561.

Bossart JL, Prowell DP (1998) Genetic estimates of population structure and gene flow: limitations, lessons, and new directions. *Trends in Ecology & Evolution*, **15**, 538–543.

Broquet T, Johnson CA, Petit E, Thompson I, Burel F, Fryxell JM (2006) Dispersal and genetic structure in the American marten, *Martes americana*. *Molecular Ecology*, **15**, 1689–1697.

Castric V, Bernatchez L (2004) Individual assignment test reveals differential restriction to dispersal between two salmonids despite no increase of genetic differences with distance. *Molecular Ecology*, **13**, 1299–1312.

Chesson P, Lee CT (2005) Families of discrete kernels for modelling dispersal. *Theoretical Population Biology*, **67**, 241–256.

Clobert J, Danchin E, Dhondt AA, Nichols JD (2001) *Dispersal*. Oxford University Press, New York.

Corbet PS (1999) *Dragonflies: Behaviour and Ecology of Odonata*. Harley Books, Colchester, UK.

DiCiccio TJ, Efron B (1996) Bootstrap confidence intervals. *Statistical Science*, **11**, 189–228.

Elliott JM (2003) A comparative study of the dispersal of 10 species of stream invertebrates. *Freshwater Biology*, **48**, 1652–1668.

Falconer DS, Mackay TFC (1996) *Introduction to Quantitative Genetics*. Longman, Harlow, UK.

Fenster CB, Vekemans X, Hardy OJ (2003) Comparison of direct and indirect estimates of gene flow in *Chamaecrista fasciculata* (Leguminosae). *Evolution*, **57**, 995–1007.

Fincke OM, Hadrys H (2001) Unpredictable offspring survivorship in the damselfly, *Megaloprepus coerulatus*, shapes parental behavior, constrains sexual selection, and challenges traditional fitness estimates. *Evolution*, **55**, 762–772.

Frankham R, Ballou JD, Briscoe DA (2002) *Introduction to Conservation Genetics*. Cambridge University Press, Cambridge, UK.

Gardarsson A, Einarsson A, Gislason GM *et al*. (2004) Population fluctuations of chironomid and simuliid Diptera at Myvatn in 1977–96. *Aquatic Ecology*, **38**, 209–217.

Goudet J (1995) *FSTAT*, version 1.2: a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.

Hanski I (2003) *Metapopulation Ecology*. Oxford University Press, Oxford.

Hansson B, Bensch S, Hasselquist D (2004) Lifetime fitness of short-and long distance dispersing great reed warblers. *Evolution*, **58**, 2546–2557.

Hardy OJ (2003) Estimation of pairwise relatedness between individuals and characterization of isolation-by-distance processes using dominant genetic markers. *Molecular Ecology*, **12**, 1577–1588.

Hardy OJ, Vekemans X (1999) Isolation by distance in a continuous population: reconciliation between spatial autocorrelation analysis and population genetic models. *Heredity*, **83**, 145–154.

Hardy OJ, Vekemans X (2002) SPAGEDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Hunger H, Röske W (2001) Short-range dispersal of the southern damselfly (*Coenagrion mercuriale*: Odonata) defined experimentally using UV fluorescent ink. *Zeitschrift Fur Okologie und Naturshutz*, **9**, 181–187.

Jolly GM (1965) Explicit estimates from capture-recapture data with both death and immigration — stochastic model. *Biometrika*, **52**, 225–247.

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

Leblois R, Estoup A, Rousset F (2003) Influence of mutational and sampling factors on the estimation of demographic parameters in a 'continuous' population under isolation by distance. *Molecular Biology and Evolution*, **20**, 491–502.

Leblois R, Rousset F, Estoup A (2004) Influence of spatial and temporal heterogeneities on the estimation of demographic parameters in a continuous population using individual microsatellite data. *Genetics*, **166**, 1081–1092.

Leblois R, Rousset F, Tikel D, Moritz C, Estoup A (2000) Absence of evidence for isolation by distance in an expanding cane toad (*Bufo marinus*) population: an individual-based analysis of microsatellite genotypes. *Molecular Ecology*, **9**, 1905–1909.

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understorey shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420–1425.

Malécot G (1948) *Les Mathematiques de l'Heredite*. Masson et Cie, Paris.

Mallet J (1986) Dispersal and gene flow in a butterfly with home range behaviour: *Heliconis erato* (Lepidoptera: Nymphalidae). *Oecologia*, **68**, 210–217.

Marr AB, Keller LF, Arcese P (2002) Heterosis and outbreeding depression in descendants of natural immigrants to an inbred population of song sparrows (*Melospiza melodia*). *Evolution*, **56**, 131–142.

Purse BV, Hopkins GW, Day KJ, Thompson DJ (2003) Dispersal characteristics and management of a rare damselfly. *Journal of Applied Ecology*, **40**, 716–728.

Purse BV, Thompson DJ (2003) Emergence of the damselflies, *Coenagrion mercuriale* (Charpentier) and *Ceriagrion tenellum* (Villers) (Odonata: Coenagrionidae), at their northern range margins, in Britain. *European Journal of Entomology*, **100**, 93–99.

Purse BV, Thompson DJ (2005) Lifetime mating success in a marginal population of a damselfly, *Coenagrion mercuriale*. *Animal Behaviour*, **69**, 1303–1315.

Raymond M, Rousset F (1995) GENEPOP, version 1.2. Population genetics software for exact tests and ecumenicisms. *Journal of Heredity*, **86**, 249–249.

Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.

Ritland K (1996) A marker-based method for inferences about quantitative inheritance in natural populations. *Evolution*, **50**, 1062–1073.

Rouquette JR, Thompson DJ (2006) Roosting site selection in the endangered damselfly, *Coenagrion mercuriale*, and implications for habitat design. *Journal of Insect Conservation*, in press.

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

Rousset F (1999) Genetic differentiation in populations with different classes of individuals. *Theoretical Population Biology*, **55**, 297–308.

Rousset F (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology*, **13**, 58–62.

Rousset F (2004) *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, Princeton, New Jersey.

Rousset F (2006) Separation of time scales, fixation probabilities and convergence to evolutionary stable states under isolation by distance. *Theoretical Population Biology*, **69**, 165–179.

Schneider C (2003) The influence of spatial scale on quantifying insect dispersal: an analysis of butterfly data. *Ecological Entomology*, **28**, 252–256.

Seber GAF (1973) *The Estimation of Animal Abundance and Related Parameters*. Griffin, London.

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

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Spong G, Creel S (2001) Deriving dispersal distances from genetic data. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **268**, 2571–2574.

Sumner J, Rousset F, Estoup A, Moritz C (2001) 'Neighbourhood' size, dispersal and density estimates in the prickly forest skink (*Gnypetoscincus queenslandie*) using individual genetic and demographic methods. *Molecular Ecology*, **10**, 1917–1927.

Vandewoestijne S, Baguette M (2004) Demographic versus genetic dispersal measures. *Population Ecology*, **46**, 281–285.

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 921–935.

Watts PC, Kemp SJ, Saccheri IJ, Thompson DJ (2005) Conservation implications of genetic variation between spatially and temporally distinct colonies of the damselfly *Coenagrion mercuriale*. *Ecological Entomology*, **30**, 541–547.

Watts PC, Rouquette JR, Saccheri IJ, Kemp SJ, Thompson DJ (2004a) Molecular and ecological evidence for small-scale isolation by distance in an endangered damselfly, *Coenagrion mercuriale*. *Molecular Ecology*, **13**, 2931–2945.

Watts PC, Saccheri IJ, Kemp SJ, Thompson DJ (2006) Impact of regional and local habitat isolation upon genetic diversity of the endangered damselfly *Coenagrion mercuriale* (Odonata: Zygoptera). *Freshwater Biology*, **51**, 193–205.

Watts PC, Thompson DJ, Kemp SJ (2004b) Cross-species amplification of microsatellite loci in some European zygopteran species (Odonata: Coenagrionidae). *International Journal of Odonatology*, **7**, 87–96.

Watts PC, Wu JH, Westgarth C, Thompson DJ, Kemp SJ (2004c) A panel of microsatellite loci for the southern damselfly, *Coenagrion mercuriale* (Odonata: Coenagrionidae). *Conservation Genetics*, **5**, 117–119.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: $F_{ST} \neq 1/(4Nm + 1)$. *Heredity*, **82**, 117–125.

Wilson AJ, Hutchings JA, Ferguson MM (2004) Dispersal in a stream dwelling salmonid: inferences from tagging and microsatellite studies. *Conservation Genetics*, **5**, 25–37.

Winters JB, Waser PM (2003) Gene dispersal and outbreeding in a philopatric mammal. *Molecular Ecology*, **12**, 2251–2259.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **28**, 114–138.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114–138.

Wright S (1946) Isolation by distance under diverse systems of mating. *Genetics*, **31**, 39–59.

This collaboration is the integration of several research projects designed to better understand the population demography and subsequent spatial genetic structure of *Coenagrion mercuriale*. Phill's research focuses on the use of genetic markers to understand determinants of spatial genetic structure and improve the management of endangered or exploited species. One of Francois's interests is to assist with the estimation of dispersal becoming a falsifiable science. Ilik is interested in the genetic determinants of fitness, and their population level consequences using quantitative/molecular genetics in conjunction with ecological field studies. Raphaël is currently working on statistical methods estimating demographic parameters from genetic data and is particularly interested in isolation by distance models and dispersal inferences. Steve is an animal geneticist whose principal focus is mapping disease resistance genes and the use of genomic technology to quantify gene expression. Dave is interested in all aspects of insect conservation and is a member of the *C. mercuriale* Biodiversity Action Plan steering group.

Appendix

Basic measures of genetic diversity across 14 microsatellite loci in 16 samples of *Coenagrion mercuriale* from the Lower Itchen Complex (LIC) and Beaulieu Heath, Hampshire, UK. *n*, sample size; H_E, expected heterozygosity; A_R, allelic richness (based on 17 individuals)

		LIC					Beaulieu Heath										
		Weh	Alm	Ivt	Ivm	Ivb	Rou	Hat	Dem	Twb	Upc	Gre	Loc	PhC	PhB	PhA	Bag
Locus	*n*	*48*	*48*	*48*	*48*	*48*	*47*	*48*	*48*	*50*	*48*	*49*	*52*	*34*	*18*	*47*	*48*
4–002	H_E	0.290	0.361	0.277	0.154	0.270	0.488	0.499	0.475	0.506	0.504	0.491	0.491	0.506	0.510	0.504	0.499
	A_R	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00
4–023	H_E	0.567	0.520	0.429	0.573	0.375	0.682	0.682	0.713	0.698	0.670	0.707	0.686	0.662	0.699	0.703	0.614
	A_R	4.00	4.00	3.99	4.96	3.99	3.84	3.75	4.41	4.17	3.00	3.93	3.89	3.00	3.94	4.62	3.73
4–024	H_E	0.526	0.512	0.456	0.456	0.497	0.538	0.514	0.621	0.573	0.650	0.504	0.625	0.593	0.611	0.630	0.502
	A_R	3.00	2.92	2.00	2.92	2.00	3.74	2.75	3.00	2.97	3.93	2.58	3.00	2.99	3.94	4.01	2.00
4–030	H_E	0.491	0.462	0.485	0.400	0.470	0.500	0.505	0.449	0.504	0.504	0.497	0.496	0.503	0.474	0.504	0.400
	A_R	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.67	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.89
4–031	H_E	0.550	0.538	0.467	0.592	0.564	0.548	0.505	0.485	0.522	0.483	0.487‡	0.517	0.365	0.595	0.571	0.623
	A_R	3.00	3.00	3.00	3.99	3.92	2.96	2.98	2.67	2.82	2.93	2.74	2.56	2.99	3.00	2.99	3.00
4–034	H_E	0.500	0.486	0.564	0.526	0.513	0.504	0.488	0.483	0.432	0.418	0.375	0.471	0.504	0.515	0.495	0.557
	A_R	2.00	2.00	3.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	3.97
4–035	H_E	0.785	0.823	0.763	0.752	0.840	0.823	0.864	0.874	0.829	0.816	0.836	0.864	0.845	0.900	0.851	0.699
	A_R	6.91	11.65	7.93	5.92	7.99	10.06	9.57	9.90	9.43	9.08	9.15	9.97	10.02	11.78	9.34	4.54
4–037	H_E	0.373	0.344	0.255	0.229	0.312	0.300	0.410	0.356	0.330	0.353	0.375	0.347	0.473	0.451	0.385	0.500
	A_R	2.00	2.96	2.00	2.00	3.00	2.95	2.75	2.00	2.89	2.59	2.73	2.70	2.99	2.94	2.00	2.00
4–042	H_E	0.325	0.257	0.321	0.284	0.368	0.502‡	0.503	0.498	0.503	0.472	0.514	0.499	0.494	0.513	0.506	0.504
	A_R	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.00	2.35	2.00	2.00	2.00	2.00	3.30
4–060	H_E	0.582†	0.636	0.612†	0.612	0.656	0.531	0.569†	0.579	0.546	0.522	0.494	0.486	0.601	0.510	0.541	0.798
	A_R	3.00	3.00	3.00	3.00	3.00	2.96	2.96	3.67	2.99	2.59	2.00	2.00	3.00	3.00	2.97	9.94
4–062	H_E	0.683	0.667	0.650	0.625	0.612	0.600	0.591	0.496	0.583	0.460	0.518	0.542	0.525	0.559	0.509	0.546
	A_R	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	2.96	3.00	2.98	3.00	3.00	2.95	3.21
4–063	H_E	0.498	0.501	0.370	0.377	0.540	0.541	0.512	0.472	0.469	0.580	0.568	0.534	0.581	0.368	0.525	0.510
	A_R	6.67	4.93	3.91	3.92	4.94	3.11	2.77	3.43	2.77	4.28	4.24	3.77	4.83	2.94	3.22	2.37
4–066	H_E	0.555	0.649	0.611	0.602	0.589	0.461	0.547	0.494	0.531	0.337	0.492	0.482	0.570	0.302	0.627	0.514
	A_R	5.94	6.00	6.91	5.92	6.92	4.45	4.11	4.89	4.76	5.38	6.06	4.92	5.06	3.94	5.94	6.65
4–067	H_E	0.664	0.778	0.736	0.771	0.737	0.791	0.861	0.819	0.856	0.851	0.798	0.791	0.805	0.828	0.835	0.836
	A_R	6.91	10.93	5.92	8.81	10.62	9.26	10.39	10.83	10.32	10.42	9.13	9.42	11.01	10.72	11.48	10.35

†indicates significant ($P < 0.05$, $k = 14$) excess of homozygotes.

‡ indicates significant ($P < 0.05$, $k = 14$) excess of heterozygotes.