MOLECULAR ECOLOGY

Effects of population characteristics and structure on estimates of effective population size in a house sparrow metapopulation

Journal:	*Molecular Ecology*
Manuscript ID:	MEC-13-1324.R3
Manuscript Type:	Original Article
Date Submitted by the Author:	23-Apr-2014
Complete List of Authors:	Baalsrud, Helle; Norwegian University of Science and Technology, Dept. of Biology; University of Oslo, Department of Biosciences; University of Oslo, Department of Biosciences Sæther, Bernt-Erik; Norwegian University of Science and Technology, Dept. of Biology Hagen, Ingerid; Norwegian University of Science and Technology, Dept. of Biology Myhre, Ane; Norwegian University of Science and Technology, Dept. of Biology Ringsby, Thor Harald; Norwegian University of Science and Technology, Dept. of Biology Pärn, Henrik; Norwegian University of Science and Technology, Dept. of Biology; Centre for Conservation Biology, Norwegian University of Science and Technology, Dept. of Biology Jensen, Henrik; Norwegian University of Science and Technology, Dept. of Biology
Keywords:	Conservation Genetics, Ecological Genetics, Population Genetics - Empirical, Birds

Title page:

Effects of population characteristics and structure on estimates of effective population size in a house sparrow metapopulation

Helle Tessand Baalsrud[1,2], Bernt-Erik Sæther[1], Ingerid Julie Hagen[1], Ane Marlene Myhre[1], Thor Harald Ringsby[1], Henrik Pärn[1] and Henrik Jensen[1]

[1] Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway.

[2] Current address: Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, NO-0316 Oslo, Norway.

Key-words: Genetic drift, dispersal, population size, sex-ratio, microsatellites, effective population size.

Author for correspondence: Henrik Jensen, Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway. E-mail: henrik.jensen@ntnu.no; fax: +47 73596100; office: +47 73596949; mobile: +47 91897064.

Running head: Effective population size in sparrows

Abstract

Effective population size (N_e) is a key parameter to understand evolutionary processes and the viability of endangered populations as it determines the rate of genetic drift and inbreeding. Low N_e can lead to inbreeding depression and reduced population adaptability. In this study we estimated contemporary N_e using genetic estimators (LDNE, ONeSAMP, MLNE and CoNe) as well as a demographic estimator in a natural insular house sparrow metapopulation. We investigated whether population characteristics (population size, sex ratio, immigration rate, variance in population size, and population growth rate) explained variation within and among populations in the ratio of effective to census population size (N_e/N_c). In general, N_e/N_c-ratios increased with immigration rates. Genetic N_e was much larger than demographic N_e, probably due to a greater effect of immigration on genetic than demographic processes in local populations. Moreover, although estimates of genetic N_e seemed to track N_c quite well, the genetic N_e estimates were often larger than N_c within populations. Estimates of genetic N_e for the metapopulation were however within the expected range ($<N_c$). Our results suggest that in fragmented populations, even low levels of gene flow may have important consequences for the interpretation of genetic estimates of N_e. Consequently, further studies are needed to understand how N_e estimated in local populations or the total metapopulation relates to actual rates of genetic drift and inbreeding.

Introduction

Effective population size (N_e) is a key parameter to understand the evolution of populations in general and the viability of small and endangered populations in particular, as it determines the rate of loss of genetic variation by genetic drift and inbreeding (Wright 1938; Frankham 1996, 2010). Furthermore, N_e affects the interplay between genetic drift and selection affecting for instance the probability of fixation of advantageous alleles (Kimura & Crow 1963). N_e is defined as the size of an ideal Wright-Fisher population in which the rate of change in heterozygosity (inbreeding effective size, N_{eI}) or allele frequencies (variance effective size, N_{eV}) is the same as in the observed population (Wright 1931). Populations with small N_e risk inbreeding depression and loss of evolutionary potential, which in turn may increase the probability of extinction (Franklin & Frankham 1998). Census population size (N_c) can be misleading in this context as $N_e << N_c$ for most natural populations (Wright 1931, 1938; Frankham 1995; Nunney 1995). Thus, knowledge about the effective population size is crucial for understanding the evolutionary processes in populations. Furthermore, for endangered populations or species, knowledge of N_e may help evaluating and minimizing any negative genetic effects. For instance, the effective population size can be maximized by artificially increasing gene-flow or carrying out strict breeding regimes (Templeton & Read 1984; Schwartz et al. 2007; Hedrick & Fredrickson 2010).

There are two different approaches to estimating N_e; using demographic ecological data or using genetic markers, or, sometimes a combination of both (Anderson & Garza 2009). The demographic approach provides an estimate of the contemporary N_e and hence current rate of genetic drift, based on demographic characteristics of the population. Unfortunately, most methods (e.g. Felsenstein 1971; Hill 1972; Engen et al. 2005) require extensive data on several demographic variables such as population size, variance in reproductive success, age-distribution, sex ratio etc. Such data are rarely obtainable for most natural populations. In addition, many of these methods are based on assumptions that are rarely fulfilled for most natural populations (e.g. stable age-distribution, constant population size, no density dependence). This is why considerable effort has been put into developing N_e estimators based on genetic data in recent years. This development has been fueled by a revolution in the advancement of techniques to efficiently genotype individuals on polymorphic molecular markers (Anderson & Garza 2009; Luikart et al. 2010).

Genetic estimates of contemporary N_e can be derived from a single sample (in time), which provides the basis for an estimate of inbreeding effective size (N_{eI}), or multiple samples spaced by one or more generations (temporal method), which can be used to estimate variance effective population size (N_{eV}) (Waples & Yokota 2007). The choice of method has several important consequences for further interpretation: N_{eI} predicts the rate at which heterozygosity is lost and depends on the number of individuals in the parent generation, whereas N_{eV} reflects the variance of change in allele frequency from one generation to the next and depends on the number of offspring (Kimura & Crow 1963; Crow & Denniston 1988; Waples 2005). As a consequence, N_{eI} will lag behind any reductions in population size or bottlenecks by at least one generation because it is affected by the population decline only when inbreeding accumulates due to increased mating between relatives (Luikart et al. 2010). In contrast, because the number of offspring usually is low in declining populations N_{eV} will decline rapidly and thus be more sensitive to reductions in population size (Luikart et al. 2010). Due to these differences N_{eV} may be more relevant when monitoring endangered species (Schwartz et al. 2007). However, N_{eI} and N_{eV} should be equal in a single isolated population of constant size (Kimura & Crow 1963).

Because of the fundamental importance of N_e in conservation, population genetics and evolutionary biology, it is relevant to quantify N_e and particularly the N_e/N_c-ratio. The N_e/N_c-ratio is affected by several factors such as sex ratio (Wright 1931), variation in family size (Wright 1938), inbreeding (Crow & Kimura 1970), age structure (Hill 1972; Engen et al. 2005), fluctuating population size (Kalinowski & Waples 2002), selection (Waples 1989); and spatial structure (e.g. Hedrick & Gilpin 1997; Wang & Caballero 1999). For more extensive reviews on N_e and N_e estimators, see Wang (2005), Anderson & Garza (2009), Charlesworth (2009), Luikart et al. (2010), and Waples (2010).

In this study we use data from a long-term study of an insular house sparrow metapopulation at Helgeland, Norway, to estimate N_e with four different genetic estimators; LDNE (Waples and Do 2008, 2010), ONeSAMP (Tallmon et al. 2008), MLNE (Wang 2001, Wang & Whitlock 2003) and CoNe (Berthier et al. 2002, Anderson 2005), and one demographic estimator (Engen et al. 2007). Our study includes 15 islands, which allows us to analyze variation in N_e/N_c locally as well as in the whole metapopulation. Estimates of several parameters known to affect N_e (Caballero 1994) such as population size (census population size, N_c), recruitment rates, adult mortality rates, sex-ratio, migration rates, and inter- and intra-individual genetic variation are available because individual-based data have been collected since 1993 (Jensen et al. 2003, 2004, 2008; Husby et al. 2006; Engen et al. 2007; Pärn et al. 2009).

We have the following objectives with this study. Initially, the congruence of different genetic estimators of N_e will be assessed by comparing estimates from different methods based on the same data set. Then for each estimator we will examine which population characteristics potentially explain variation in the N_e/N_c–ratio across local populations and years within this metapopulation. Furthermore, we will compare N_e estimates from analyses of genetic data with demographic estimates of N_e (Engen et al. 2007). Finally, we will investigate how population structure and gene-flow affect genetic N_e estimates by comparing N_e calculated for the entire metapopulation ($metaN_e$) with the sum of N_c for local populations (metaN_c) under the prediction that the $metaN_e/metaN_c$-ratio > 1 in Wright's island model (Wang & Caballero 1999). We will also compare $metaN_e$ with the sum of N_e for all local populations (ΣN_e), predicting that given a Wright's island model $metaN_e > \Sigma N_e$, but given more complex metapopulation dynamics this relationship will be the opposite ($metaN_e < \Sigma N_e$) (Hedrick & Gilpin 1997;Wang & Caballero 1999).

Materials and Methods

Study system

The study area consisted of fifteen islands in Northern Norway from Sleneset (66°22´N, 12°36´E) in the southeast to Myken (66°46´N, 12°29´E) in the northwest (Fig. 1). These islands were populated by house sparrows continuously or periodically during the 17 year study period (1993-2009). For more extensive information regarding this study system, see Sæther et al. (1999), Ringsby et al. (2002), Pärn et al. (2009, 2012).

Data collection and sampling scheme

From 1993-2009 we captured adult and juvenile individuals using mist nets, while nestlings were sampled from the nest. A blood sample (25 μL by venipuncture) was collected from each individual. We designated each bird a metal ring with an individual number and a unique combination of three colored bands on its tarsi. This allowed us to estimate various demographic parameters from recapture and observation data. As the average generation time for the house sparrow is approximately 2 years (Jensen et al. 2008) we assumed that samples spaced by 3 years were from separate generations. To represent six generations we selected the following years: 1994, 1997, 2000, 2003, 2006 and 2009. For the single sample estimators of N_e, we obtained one point estimate for each population in each of these years. For the temporal estimators of N_e, we used data from pairs of samples spaced by 1, 4 or 7 generations (i.e. 3, 9 or 15 years respectively) for each population.

Population characteristics

We estimated annual adult census population size (N_c) in one of two ways: on the islands where the percentage of marked individuals was high (>70%, and usually close to 100%) we estimated N_c as the number of marked adult individuals that were either captured or observed in a given year, or captured/observed in both a previous and a subsequent year (Jensen et al. 2006, 2013). Otherwise we estimated annual N_c by counting number of adults present in the population at the start of the breeding season (Pärn et al. 2012). There is a strong correlation between these two methods ($r=0.959$, $P<0.001$; see Jensen et al. 2013), and N_c was assumed equal to actual adult population size. For single sample estimators we compared N_c with $\widehat{N}_e$, but for temporal estimators we compared $\widehat{N}_e$ with the harmonic mean census population size (N_H) across the years since the previous sampling event (both years of sampling included). This is because the single sample $\widehat{N}_e$ represents N_e at the time of sampling, whereas the temporal $\widehat{N}_e$ represents the harmonic mean N_e in the time interval considered (Waples 2010). We calculated the variance in population size ($\sigma^2_{N_c}$) and the population growth rate (dN/dt) between two samples; for single sample $\widehat{N}_e$ between the sampling year and the previous sampling year, for temporal $\widehat{N}_e$ between the two sampling years. The sex ratio (SR) was defined as the proportion of males in the population (for temporal $\widehat{N}_e$ the SR was averaged over the two sampling years). We calculated the average immigration rate (m) for either the generation preceding the sampling event (single sample $\widehat{N}_e$) or the time interval between sampling years (temporal $\widehat{N}_e$). Information on sampling schemes can be found in Supporting Information (Table S3 (single sample) and Table S4 (temporal)).

Molecular analyses

We extracted DNA from blood samples as described in Elphinstone et al. (2003) and used polymerase chain reactions (PCR) to amplify DNA and genotype each individual on 13 polymorphic nuclear microsatellite loci (see Supporting Information Appendix S1) that appeared to be selectively neutral, unlinked and in Hardy Weinberg equilibrium. For more information on the genotyping and use of these loci in e.g. population genetic studies see Griffith et al. (2007), Kekkonen et al. (2011), Billing et al. (2012), Dawson et al. (2012) and Jensen et al. (2013). The widespread and successful use of these loci suggests they are suitable for estimating genetic effective population size.

Estimation of genetic N_e

Single sample estimators

LDNE

The LDNE program implements a moment-based method for estimating N_e based on linkage disequilibrium (LD), defined as the non-random association of alleles at different loci, that arises due to random genetic drift (Waples & Do 2008, 2010). Although single sample estimates of N_e usually apply to the parental generation, estimates based on LD may provide information on N_e more than one generation prior to the sampling because LD may take several generations to decay, particularly if linked loci are used (Waples 2010, Luikart et al. 2010). Here we assumed random mating, as the house sparrows in this study system appears not to be strictly monogamous (Jensen et al. 2008).

ONeSAMP

The ONeSAMP program implements approximate Bayesian computation (ABC) to estimate N_e by comparing eight summary statistics (including LD) obtained for the population in question with the same statistics calculated for 50 000 simulated populations (Tallmon et al. 2008). We chose two as the lower bound of the prior, and because N_e theoretically can be at most twice as high as N_c (Wright 1938), $2N_c$ was chosen as its upper bound. The repeat motif was specified for each locus (for repeat motif for the different loci, see Griffith et al. (2007) and Dawson et al. (2012)).

Temporal methods: Multiple samples estimators

For both temporal methods the upper bound of the prior in the estimation procedure was chosen to be $2N_c$ for the sampling year with the highest N_c.

MLNE

The MLNE method estimates $\widehat{N}_e$ from temporally spaced samples using a pseudo-likelihood method which assumes that temporal changes in allele frequencies are caused by genetic drift alone (Wang 2001; Wang & Whitlock 2003). The estimation procedure is based on the Wright-Fisher model, but has less restrictive assumptions as it allows for migration (open populations) assumed to be from an infinite, unchanging source population, estimating $\widehat{N}_e$ and $\widehat{m}$ jointly (Wang & Whitlock 2003).

We estimated $\widehat{N}_e$ assuming both isolated populations ($\widehat{N}_{e(MLNE,\ closed)}$), and open populations ($\widehat{N}_{e(MLNE,\ open)}$). Consequently, $\widehat{N}_{e(MLNE,\ open)}$ are estimates for which a “genetic immigration rate” (referred to as ($\widehat{m}_{(MLNE)}$) has been taken into account. When estimating $\widehat{N}_{e(MLNE,\ open)}$ we pooled individuals from all the island populations except the focal population, and defined this pool as the source population for any migrants. Furthermore, MLNE requires that a relationship between drift and migration is specified; we assumed non-equilibrium as this is more realistic for the small populations in this study.

To examine whether choice of prior affected our results we also estimated N_e using the MLNE(closed) method when the upper bound of the prior was set to $40N_c$

for the sampling year with the highest N_c (see Supporting Information Tables S2 and S4).

CoNe

The program CoNe gives the likelihood of N_e given genetic data sampled from the same population at different points in time (Anderson 2005). This method is based on coalescent theory and assumes that coalescent events are only driven by genetic drift while ignoring mutations and dispersal (Anderson 2005).

Metapopulation N_e

Genetic data from each local population was pooled for each year or combination of years to estimate *meta*N_e for each estimator. Additionally, N_e-values were summed across islands for each year/combinations of years, giving ΣN_e. Census population size was summed across islands to give an estimate of *meta*N_c. When N_e for a local population was not available (see Supporting Information Table S2) the same population was excluded from the other estimates to make direct comparisons possible. Note that neither the *meta*N_e nor the ΣN_e are expected to be entirely correct estimates of N_e for a metapopulation. We therefore only qualitatively compared estimates of *meta*N_e and ΣN_e with *meta*N_c, and *meta*N_e/*meta*N_c–ratios with subpopulation N_e/N_c-ratios to examine variation in N_e at different population levels, and thus indicate how ignoring population structure may affect estimates of N_e.

Estimation of demographic N_e

Demographic methods are often based on very restrictive assumptions (e.g. Felsenstein 1971; Hill 1972) or a large number of parameters (e.g. Engen et al. 2005). Here we based our estimates on the approach of Engen et al. (2007), which assumes constant mean vital rates independent of age. This simplifies the estimation considerably. N_e was calculated for each sex separately as vital rates can be sex-specific. N_e for females (N_{ef}) was based on a simplification of Engen et al. (2005) and is given by

$$N_{ef} = \frac{N_f}{\sigma^2_{dgf} T_f} = \frac{N_f}{\left[b_f/4 + \sigma^2_f/4 + s_f\left(1 - s_f\right) + c_f\right] T_f} \qquad (2)$$

where N_f is the number of females, σ^2_{dgf} is the demographic variance of a hypothetical female subpopulation of heterozygotes carrying a rare allele, b_f is the mean number of female offspring born to each female, σ^2_f is the variance in number of female offspring per female, s_f is the probability of survival for females, c_f is the covariance between an individual's number of offspring and the indicator variable (0 or 1) for its survival, and T_f is the generation time for the female population given by $T_f = \lambda/(\lambda - s_f)$ where λ is the deterministic growth rate. N_e for males (N_{em}) was calculated in the same way. Second, the N_e of the total population was calculated as follows, based on Wright's formula for uneven sex ratios modified to also allow for non-overlapping generations (Engen et al. 2007)

$$N_e = \frac{4\lambda^2 b N_{ef} N_{em}}{b_f N_{ef} + b_m N_{em}} \qquad (3)$$

$\widehat{N}_{e(demographic)}$ was obtained from the same sampling intervals as temporal genetic $\widehat{N}_e$, by multiplying the mean population size during the sampling interval for each island with the N_e/N_c-ratio given in Table 1 in Engen et al. (2007). Note that this approach assumes a constant N_e/N_c-ratio across years. Metapopulation N_e was not estimated using the demographic method as this approach was currently only possible for a limited number of years on 6 out of 15 islands. More information about the method and exact values used in the calculations can be found in Engen et al. (2007).

Statistical analyses

We used Pearson's correlation coefficient (r) to investigate the relationship between estimates from different estimators ($\widehat{N}_{e(LDNE)}$ and $\widehat{N}_{e(ONeSAMP)}$, and $\widehat{N}_{e(MLNE)}$ and $\widehat{N}_{e(CoNe)}$), and between $\widehat{N}_e$ and N_c (N_H for temporal estimates), and finally between $\widehat{N}_e$ and $\widehat{N}_{e(demographic)}$. In order to determine the importance of population characteristics for $\widehat{N}_e/N_c$, we modeled $\widehat{N}_e/N_c$ and $\widehat{N}_e/N_H$ as a function of the following predictor variables: SR, m, N_c or N_H, dN/dt, $\sigma^2_{N_c}$ and the number of generations between samples (g). Additionally, the interactions were included in the *a priori* global models: $N_c \times SR$ and $N_c \times m$ as the effect of SR and m on $\widehat{N}_e/N_c$ could vary with N_c (N_H for temporal estimators). Hence, the global models included five or six covariates (for single sample and temporal methods, respectively) and two two-way interactions. All 52 or 104 models nested within the global models were tested (for single sample and temporal methods, respectively). Analyses were carried out using both generalized linear models (GLM) with a Gaussian error structure (using the lm function in R; R Development Core Team 2011) and generalized linear mixed models (GLMM, with a Gaussian error structure, using the nlme package (Pinheiro et al. 2011)) with population as a random factor. As GLMs and GLMMs gave similar results only the results from the GLMs are presented. Model selection was carried out using Akaike's

Information Criterion with a correction for smaller sample sizes (AIC_C) following Burnham & Anderson (2002). Detailed results of the model selection procedures are given in Supporting Information Appendix S2. All statistical analyses were carried out using R (R Development Core Team 2011). For an overview of removed data, including justification for removing outliers, see Supporting Information Table S2.

Results

There was large variation in N_c within and among insular house sparrow populations in Northern Norway across the 17 year study period (Fig. 2, 3). Aldra was colonized in 1998 and populated continuously thereafter (see Billing et al. 2012). The populations on Sundøy and Ytre Kvarøy went extinct in 2000 (see Ringsby et al. 2006), and the Selvær population went effectively extinct in 2000 (only four males present), but quickly rebounded due to immigration (see Supporting Information Figure S2). For the other island populations the population sizes ranged from less than 10 individuals (e.g. Selsøyvik) to more than 150 individuals (e.g. Hestmannøy).

Estimates of local N_e

$\widehat{N}_{e(LDNE)}$ and $\widehat{N}_{e(ONeSAMP)}$ were both significantly positively correlated with N_c and tracked fluctuations in N_c over time (Fig. 2, Table 1, Supporting Information Table S3). The two estimators showed different patterns; $\widehat{N}_{e(LDNE)}$ were mostly lower than N_c in relatively large populations (i.e. populations larger than $N_c \approx 25$, see Supporting Information Fig. S1), while $\widehat{N}_{e(ONeSAMP)}$ typically were higher than N_c. For small populations, the relationship between $\widehat{N}_e$ and N_c appeared to be opposite for both estimators (Supporting Information Fig. S1).

$\widehat{N}_{e(MLNE,\ closed)}$ and $\widehat{N}_{e(MLNE,\ open)}$ were significantly positively correlated with each other and with N_H (Table 1), whereas $\widehat{N}_{e(CoNe)}$ was significantly correlated only with $\widehat{N}_{e(MLNE,\ open)}$ (Table 1). All three temporal estimators seemed to track fluctuations in N_H over time (Fig. 3, Supporting Information Table S4). Temporal $\widehat{N}_e$ was generally higher than N_H; except for $\widehat{N}_{e(MLNE,\ open)}$ (Fig. 3). Overall the MLNE and CoNe methods produced quite similar estimates, although the estimates given by CoNe were generally higher than those from MLNE (Fig. 3).

Estimates of immigration rates from MLNE ($\widehat{m}_{(MLNE)}$) ranged from 0.00 to 1.00, with a mean value of 0.43 (Supporting Information Table S4). These estimates were much higher than the observed (ecological) migration rate m calculated based on observed natal dispersal events of recruiting individuals between the islands (range: 0.00 - 0.14, mean = 0.04; see Pärn et al. 2009, 2012 and Supporting Information Figure S2).

Population characteristics and variation in local N_e/N_c

The most parsimonious model explaining 35% of the variation in $\widehat{N}_{e(LDNE)}/N_c$ included two parameters: sex ratio (SR) and immigration rate (m) (model 1, Table 2a). SR and m were positively related to $\widehat{N}_{e(LDNE)}/N_c$ (Table 3a), indicating that $\widehat{N}_{e(LDNE)}$ was relatively higher compared to N_c when the population was more male biased and there were more immigrants.

The most parsimonious model explaining 31% of the variation in $\widehat{N}_{e(ONeSAMP)}/N_c$ included only N_c (model 1, Table 2b). This model showed that N_c was positively related to $\widehat{N}_{e(ONeSAMP)}/N_c$ (Table 3b), thus $\widehat{N}_{e(ONeSAMP)}$ was relatively higher compared to N_c at higher values of N_c (see also Supporting Information Fig. S1b).

Variation in $\widehat{N}_{e(MLNE,\ closed)}/N_H$ was best explained by population size (N_H), sex ratio (SR), immigration rate (m) and population growth rate (dN/dt) (model 1, Table 2c). In this model, which explained 35% of the variance in $\widehat{N}_{e(MLNE,\ closed)}/N_H$, $\widehat{N}_{e(MLNE,\ closed)}/N_H$ was negatively related to N_H and SR, and positively related to m and dN/dt (Table 3c). This implied that $\widehat{N}_{e(MLNE,\ closed)}$ was relatively smaller compared to

N_H at higher population sizes and at higher proportions of males in the population, and relatively higher than N_H at higher population growth rates and immigration rates.

According to the best model, which explained 47% of the variation in $\widehat{N}_{e(MLNE,\ open)}/N_H$ (model 1, Table 2d), N_H was negatively related to $\widehat{N}_{e(MLNE,\ open)}/N_H$ and $\sigma^2_{N_c}$ were positively associated with $\widehat{N}_{e(MLNE,\ open)}/N_H$ (Table 3d). Consequently, $\widehat{N}_{e(MLNE,\ open)}$ was relatively larger compared to N_H at smaller values of N_H and with higher $\sigma^2_{N_c}$.

The most parsimonious model explaining 71% of the variation in $\widehat{N}_{e(CoNe)}/N_H$ included population size (N_H), immigration rate (m), population growth rate (dN/dt) and the interaction term $N_H \times m$ (model 1, Table 2e). Parameter estimates for model 1 (Table 3e) showed that m and dN/dt had a positive effect on $\widehat{N}_{e(CoNe)}/N_H$, whereas N_H and the interaction term $N_H \times m$ had negative parameter estimates. However, the main effect of N_H was not significant ($p = 0.20$). Thus, $\widehat{N}_{e(CoNe)}$ was relatively higher compared to N_H when the number of immigrants increased and with higher population growth rate. The magnitude of the positive effect of m on $\widehat{N}_{e(CoNe)}/N_H$ was reduced when population size increased.

The relationship between genetic and demographic local N_e

The estimates of N_e from the demographic method were significantly positively correlated with estimates from the MLNE method, as well as with estimates from the CoNe method (Table 1). Thus, this suggests that these methods reflected current rates of drift in these populations. However, both the MLNE method and the CoNe method always produced estimates that were larger than $\widehat{N}_{e(demographic)}$ (see Supporting Information Figure S1).

The effect of population structure

Metapopulation estimates of N_e based on the pooled samples from all local populations for a given year (or two points in time for the temporal estimator) were only estimated for ONeSAMP, LDNE and MLNE (closed). Metapopulation N_e could not be estimated using CoNe due to too many missing estimates for each local population (see Supporting Information Table S2) and estimating metapopulation N_e using MLNE(open) is not possible as we currently do not have data on the genetic composition of the potential source of immigrants into the house sparrow metapopulation. Even though the sample size was too small for proper statistical testing, some patterns were apparent; for LDNE, ΣN_e was lower than metaN_c, while for MLNE it was always higher (Figure 4, Supporting Information Table S5). For both LDNE and MLNE ΣN_e was in general more than twice the $metaN_e$ (Figure 4, Supporting Information Table S5). Also, for both LDNE and MLNE $metaN_e/metaN$ was in the range 0.260-2.521, and mostly below 1 (Supporting Information Table S5).

Discussion

As expected, estimates of local N_e within island populations were strongly positively related to N_c (Table 1, Fig. 2 and 3). However, although estimates of genetic N_e seemed to track N_c quite well, N_e estimates within local populations were in general larger than census population size (Fig. 2 and 3), with $\widehat{N}_{e(LDNE)}$ being the only exception (Fig. 2). On the metapopulation level we found that $metaN_e$ was usually smaller than $metaN_c$ (Fig. 4).

There was an overall congruence between different genetic estimators of N_e, and between temporal genetic and demographic estimators of N_e (Table 1). Temporal genetic N_e was however always larger than demographic N_e. The differences between the estimators are partly due to the estimators being affected differently by immigration and fluctuating population size (see discussion below). The estimators can also be inaccurate, which is very likely in many natural populations given the restrictive assumptions underlying the estimators (constant population size, no gene-flow, non-overlapping generations etc.). As we do not know the true value of N_e it is however difficult to assess the magnitude and direction of a potential bias. For example, when the upper bound of the prior was increased from 2 times N_c to 40 times N_c many of the N_e-estimates from the MLNE(closed) method increased considerably, showing that this estimator is sensitive to the prior distribution (Supporting Information Table S4). The N_e-estimates increased on average by more than 10-fold when the two population samples were spaced only one generation apart but remained almost the same when spaced four or seven generations apart (Supporting Information Fig. S3). This suggests that the bias introduced by age structure when using a method which assumes non-overlapping generations (such as e.g. MLNE) can be reduced by increasing the number of generations between the samples (Waples & Yokota 2007). Furthermore, some estimates were either infinite or clear outliers (see Supporting Information Tables S2, S3 and S4), perhaps due to the combined effects of relatively few genetic marker loci and small sample sizes (Waples 1989; England et al. 2006). Although these estimates were excluded from the analyses, they do highlight a general concern with at least some of the genetic N_e estimators; $\hat{N}_e$ may in certain cases be very biased. Given that $N_e > N_c$ for most estimates in our study (Table 1, Fig. 2 and 3) it seems likely that there exist an upward bias for genetic N_e estimators on the local population level. The most likely cause of this potential bias is immigration. We recommend the use of different estimators of N_e to obtain some notion of how robust the estimate is, and that $\hat{N}_e$ should be interpreted with caution if no other information is available from the population in question. Different bounds of the prior should also be tried out; if the estimates returned are much higher and even equal to the upper bound of the prior this might indicate that there is not enough information in the molecular data for the method to properly quantify the genetic drift component and give reliable estimates of N_e, perhaps because too few generations have passed between the temporal samples.

Population characteristics and variation in N_e/N_c

Contemporary estimates of N_e will mainly reflect local demographic and evolutionary processes that have occurred during recent generations(Waples 2010), with single sample estimators generally reflecting processes that occurred in the parental generation, and temporal estimators reflecting the processes that have occurred during the time span considered (Luikart et al. 2010). Accordingly, demographic characteristics of the populations during one or a few generations prior to sampling explained between 31 and 71 % of the observed variance in local N_e/N_c for the different genetic estimators. This pattern was however complex, as the population characteristics affecting the N_e/N_c-ratio differed between estimators (Table 2 and 3).

Population size and population growth rate

According to theory there should be no relationship between N_e/N_c and N_c (Kalinowski & Waples 2002). However, N_c positively affected N_e/N_c for the ONeSAMP method and negatively affected N_e/N_H for the MLNE and CoNe methods (Table 3). A negative relationship between N_e/N_c and N_c was found in other studies as well (Ardren & Kapuscinski 2003; Beebee 2009). These studies attributed this pattern to either genetic compensation (which is a higher than expected N_e at low values of N_c because reproductive variance may be lower in small populations), or simply an artifact of plotting a fraction against its denominator. ONeSAMP gave relatively higher N_e with higher N_c prior, consistent with results from Phillipsen et al. (2011), and *meta*N_e estimates for this estimator were therefore extremely high (i.e. >2500). This is most likely an artifact of the estimation procedure used in ONeSAMP, which seems to be inappropriate for higher values of N_e as the signal from drift attenuates with increasing N_e (see Phillipsen et al. (2011)), and not an effect of population structure. As expected from theory, N_e/N_H was positively related to population growth rate for the two temporal methods, which estimate N_{eV} (Waples 2005). Fluctuations in population size are expected to influence N_{eI} and N_{eV} differently, as variance in allele frequencies and inbreeding relates differently to population dynamics (Crow & Denniston 1988). Changes in N_{eV} are expected to follow changes in N_c because variance in allele frequency is directly dependent on N_c. On the other hand, N_{eI} will lag by at least one generation, as it relates to the number of parents that produced the sample. Although it may seem that the single sample estimators track N_c better than the temporal estimators (Fig. 2, 3), there is no evidence that this is the case based on the observed correlations between N_e estimates and N_c (Table 1).

Variance in population size should lead to a reduction in N_e and N_e/N_c (Kalinowski & Waples 2002). Therefore, the positive relationship between $\widehat{N}_{e(MLNE,open)}$ and $\sigma^2_{N_c}$ is puzzling. This relationship could however arise if there is a positive relationship between population size and the magnitude of fluctuations in population size.

Sex ratio

A prediction from Wright's theory is that a skewed sex ratio will decrease N_e toward the effective size for the rarest sex (Wright 1931, 1938). Empirical studies have reported that unequal sex ratio had a negative effect on N_e/N_c (Frankham 1995). However, the effect of sex ratio on N_e/N_c depends on the mating system (Nunney 1993). The house sparrow mating system is probably dominance polygyny, resulting in higher variance in reproductive success for males than females (Anderson 2006; but see Jensen et al. 2004). The maximum value of N_e/N_c is then expected in a male biased population (Nunney 1993). This is concordant with the results for the LDNE method, where sex ratio (i.e. proportion of males) affected N_e/N_c positively (Table 3a). However, for MLNE(closed) the opposite result was found (Table 3c). This could be due to effects of the other parameters included in the model (population size, immigration and population growth rate), as other factors could interact with sex ratio and influence its relationship with N_e/N_c.

Immigration

Immigration rate had a positive effect on N_e/N_c for three of the estimators: LDNE, MLNE(closed) and CoNe (Table 3). The positive relationship between immigration rate and N_e/N could be a direct consequence of the increased genetic variation introduced by immigrants (Charlesworth 2009), because immigrants that are genetically different from residents are expected to increase levels of LD and hence reduce N_e in the local population (Waples & England 2011). The effect of immigration will therefore depend on the genetic differentiation between the source population and recipient population (Wang & Whitlock 2003). Preliminary analyses suggest that the average pairwise F_{ST} value among the island populations in this metapopulation is ca. 0.03, indicating moderate genetic differentiation (Jensen et al. in prep.). Moreover, Jensen et al. (2013) found that genetic differentiation between island populations increased with distance on a larger scale along the Norwegian coast, and previous studies have shown that dispersal distances in the house sparrow are generally very short (Tufto et al. 2005; Anderson 2006) and occur most frequently between neighboring islands (Pärn et al. 2012). As a consequence, the average immigrant seems likely to be sufficiently genetically similar to the average individual in the recipient population to avoid increasing LD levels at the typed genetic markers, and the direct effect of immigration to reduce the rate of loss of genetic variation is probably the cause for its positive effect on N_e/N_c. This assumes that the recorded migrants are a part of the breeding population, which is not always the case in our study metapopulation in which male immigrants have lower fitness than resident males (Pärn et al. 2009). A comparison of the temporal genetic estimates from MLNE and CoNe and the demographic estimates of N_e revealed that demographic estimates were not affected by immigration in the same way, as the values of local N_e fall within what is expected with respect to N_H ($N_e < N_H$). The demographic method is on the other hand sensitive to local population demography such as the mean and variance in vital rates and deviation from a 1:1 sex ratio (Engen et al. 2005, 2007). Thus, immigration is expected to reduce demographic N_e slightly because the lowered fitness of male immigrants would reduce the mean and increase the variance in male reproductive success (see eqn. 2). In contrast, genetic N_e will be directly affected and elevated by immigration, because immigration counteracts the effect of genetic drift. The true N_e probably lies between N_e estimated using the demographic approach and N_e estimated using genetic methods.

Immigration rates estimated using the MLNE method ($\widehat{m}_{(MLNE)}$, mean = 0.43) were much higher than what is reasonable given the (ecological) m (mean = 0.044), calculated based on observed dispersal events (Pärn et al. 2009, 2012; see also Supporting Information Figure S2). The reason for the improbably high $\widehat{m}_{(MLNE)}$ may be related to the fact that individuals from all other islands than the one for which N_e was estimated were pooled and defined as the source population in our analyses. Hence, a continent-island system was assumed, when in fact the metapopulation is an island-island system, with local populations of very different sizes and asymmetrical exchange of migrants (Pärn et al. 2012, Supporting Information Figure S2). Similarly, improbably large estimates of m were also found in the frog *Rana pipiens* and were also attributed to the MLNE method's assumption regarding source of immigrants (Hoffman et al. 2004).

Population structure and the importance of spatial scale

Given that gene-flow between local populations has such a large impact on N_e it may be more appropriate to study N_e for the total metapopulation ($metaN_e$). However, this may be very challenging as the metapopulation N_e is not equal to the sum of all the local population N_e-values (Hedrick & Gilpin 1997). Under Wright's island model, N_e for a metapopulation is higher than the equivalent panmictic population. However, many assumptions of the island model are unrealistic in natural populations. For example, if there is higher variance in reproductive success between local populations than expected by a Poisson distribution or if extinction-recolonization dynamics are accounted for, then metapopulation N_e will be (much) smaller than N_e for a panmictic population with the same N_c (Hedrick & Gilpin 1997, Whitlock & Barton 1997, Wang & Caballero 1999, Ovaskainen & Hanski 2004). In our house sparrow study metapopulation we have shown that $metaN_e$ generally is smaller than the metapopulation census population size ($metaN_c$) and clearly smaller than the sum of the local population N_e-values (ΣN_e; Fig. 4, Supporting Information Table S5). The true value of N_e for the total metapopulation probably lies in the range between $metaN_e$ and $metaN_c$. Proper estimation of genetic metapopulation N_e using e.g. the model of Whitlock & Barton (1997) was not possible in this study, but should be carried out when information on the variance among local populations in reproductive success is available and can be combined with information on local population sizes and the level of genetic differentiation among populations (measured by Wright's F_{ST}, which depends on dispersal rates).

Our results strongly show the importance of identifying the proper spatial scale for estimating N_e; if unaware of population structure and metapopulation dynamics one might risk either overestimating local N_e or underestimating metapopulation N_e (see also Fraser et al. 2007, Palstra & Ruzzante 2011). Our study also suggests the importance a metapopulation structure can have for preserving genetic variation, especially when the environmental stochasticity is high. The different island populations have undergone major fluctuations in population size, including severe bottlenecks and extinction events (Fig. 2 and 3). However, LDNE $metaN_e$ (but not single-generation MLNE $metaN_e$) remains relatively stable across years (Fig. 4). One can speculate whether this may be due to the stabilizing effect of migration between islands; we have shown here that migration has a positive impact on local N_e and thus maintenance of genetic variation.

Conclusions and implications

Genetic N_e estimated with proper caution with respect to potential biases or imprecisions, can be used to guide management decisions (Leberg 2005). Identifying N_e and the factors causing low values of N_e is vital for conservation (Wang 2009). Knowledge of how N_e can be maximized by management is also necessary. From our analyses it is clear that population size itself is an important factor, as N_e increased with N_c for all estimators. Frankham (1995) conclude that a fluctuating population size is the most important factor responsible for reducing the N_e/N_c ratio. Immigration had a positive effect on N_e/N_c and facilitating gene flow in fragmented habitats may therefore be an important conservation measure to reduce loss of genetic variation. We suggest that more effort should be put into providing empirical estimates of N_e for both local populations and metapopulations. Because more populations will become fragmented in the future, an increased understanding of how different factors affect the rate of genetic drift at every level in such systems is essential (Hedrick & Gilpin 1997, Waples 2010).

Acknowledgements

We thank the inhabitants in our study area whose hospitality and friendliness made this study possible. We are also grateful to our many field workers for assistance in collecting the data, Atle Torvik Kristiansen for help with the genotyping, Ivar Herfindal and Thomas Kvalnes for help with R, and four anonymous referees for helpful comments on previous versions of the paper. This work was supported by grants from the European Research Council (ERC-2010-AdG 268562), the Research Council of Norway (FRIBIO 204303 and 221956), and NTNU. The research was carried out in accordance with permits from the Norwegian Environment Agency and the Bird Ringing Centre at Stavanger Museum, Norway.

Literature cited

Anderson EC (2005) An efficient Monte Carlo method for estimating Ne from temporally spaced samples using a coalescent-based likelihood. Genetics, **170**, 955-967.

Anderson EC, Garza JC (2009) Estimation of population size with molecular genetic data, Technical Memorandum NMFS. NOAA, California.

Anderson TR (2006) Biology of the ubiquitous house sparrow - From genes to populations. Oxford University Press, New York.

Ardren WR, Kapuscinski AR (2003) Demographic and genetic estimates of effective population size (Ne) reveals genetic compensation in steelhead trout. Molecular Ecology, **12**, 35-49.

Beaumont MA (2003) Estimation of population growth or decline in genetically monitored populations. Genetics, **164**, 1139-1160.

Beebee TJC (2009) A comparison of single-sample effective size estimators using empirical toad (*Bufo calamita*) population data: genetic compensation and population size-genetic diversity correlations. Molecular Ecology, **18**, 4790-4797.

Berthier P, Beaumont MA, Cornuet JM, Luikart G (2002) Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: A genealogical approach. Genetics, **160**, 741-751.

Billing AM, Lee AM, Skjelseth S, Borg AA, Hale MC, Slate J, Parn H, Ringsby TH, Sæther B-E, Jensen H (2012) Evidence of inbreeding depression but not inbreeding avoidance in a natural house sparrow population. Molecular Ecology, **21**, 1487-1499.

Burnham KP, Anderson DR (2002) Model selection and multimodal inference: A practical information-theoretic approach. 2nd edition. Springer Verlag, New York.

Caballero A (1994) Developments in the prediction of effective population size. Heredity, **73**, 657-679.

Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. Nature Reviews Genetics, **10**, 195-205.

Crow JF, Denniston C (1988) Inbreeding and variance effective population numbers. Evolution, **42**, 482-495.

Dawson DA, Horsburgh GJ, Krupa AP, Stewart IRK, Skjelseth S, Jensen H, Ball AD, Spurgin LG, Mannarelli ME, Nakagawa S, Schroeder J, Vangestel C, Hinten GN, Burke T (2012) Microsatellite resources for Passeridae species: a predicted microsatellite map of the house sparrow Passer domesticus. Molecular Ecology Resources, **12**, 501-523.

Elphinstone MS, Hinten GN, Anderson MJ, Nock CJ (2003) An inexpensive and high-throughput procedure to extract and purify total genomic DNA for population studies. Molecular Ecology Notes, **3**, 317-320.

Engen S, Lande R, Sæther B-E (2005) Effective size of a fluctuating age-structured population. Genetics, **170**, 941-954.

Engen S, Ringsby TH, Sæther B-E, Lande R, Jensen H, Lillegard M, Ellegren H (2007) Effective size of fluctuating populations with two sexes and overlapping generations. Evolution, **61**, 1873-1885.

England PR, Cornuet JM, Berthier P, Tallmon DA, Luikart G (2006) Estimating effective population size from linkage disequilibrium: severe bias in small samples. Conservation Genetics, **7**, 303-308.Felsenstein J (1971) Inbreeding and variance effective numbers in populations with overlapping generations. Genetics, **68**, 581-597.

Frankham R (1995) Effective population size / adult population size ratios in wildlife: a review. Genetical Research, **66**, 95-107.

Frankham R (1996) Relationship of genetic variation to population size in wildlife. Conservation Biology, **10**, 1500-1508.

Frankham R (2005) Genetics and extinction. Biological Conservation, **126**, 131-140.

Frankham R (2010) Challenges and opportunities of genetic approaches to biological conservation. Biological Conservation, **143**, 1919-1927.

Franklin IR, Frankham R (1998) How large must populations be to retain evolutionary potential? Animal Conservation, **1**, 69-70.

Fraser DJ, Hansen MM, Ostergaard S, Tessier N, Legault M, Bernatchez L (2007) Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. Molecular Ecology, **16**, 3866-3889.

Griffith SC, Dawson DA, Jensen H, Ockendon N, Greig C, Neumann K, Burke T (2007) Fourteen polymorphic microsatellite loci characterized in the house sparrow *Passer domesticus* (Passeridae, Aves). Molecular Ecology Notes, **7**, 333-336.

Hedrick PW, Fredrickson R (2010) Genetic rescue guidelines with examples from Mexican wolves and Florida panthers. Conservation Genetics, **11**, 615-626.

Hedrick PW, Gilpin ME (1997) Genetic effective size of a metapopulation.*in* I. A. Hanski and M. E. Gilpin, editors. Metapopulation biology - Ecology, genetics, and evolution. Academic Press, London.

Hill WG (1972) Effective size of populations with overlapping generations. Theoretical Population Biology **3**:278-&.

Hill WG (1981) Estimation of effective population size from data on linkage disequilibrium. Genetical Research, **38**, 209-216.

Husby A, Sæther B-E, Jensen H, Ringsby TH (2006) Causes and consequences of adaptive seasonal sex ratio variation in house sparrows. Journal of Animal Ecology, **75**, 1128-1139.

Jensen H, Moe R, Hagen IJ, Holand AM, Kekkonen J, Tufto J, Sæther B-E (2013) Genetic variation and structure of house sparrow populations: is there an island effect? Molecular Ecology, **22**, 1792-1805.

Jensen H, Steinsland I, Ringsby TH, Sæther B-E (2008) Evolutionary dynamics of a sexual ornament in the house sparrow (*Passer domesticus*): The role of indirect selection within and between sexes. Evolution, **62**, 1275-1293.

Jensen H, Svorkmo-Lundberg T, Ringsby TH, Sæther B-E (2006) Environmental influence and cohort effects in a sexual ornament in the house sparrow, *Passer domesticus*. Oikos, **114**, 212-224.

Jensen H, Sæther B-E, Ringsby TH, Tufto J, Griffith SC, Ellegren H (2003) Sexual variation in heritability and genetic correlations of morphological traits in house sparrow (*Passer domesticus*). Journal of Evolutionary Biology, **16**, 1296-1307.

Jensen H, Sæther B-E, Ringsby TH, Tufto J, Griffith SC, Ellegren H (2004) Lifetime reproductive success in relation to morphology in the house sparrow *Passer domesticus*. Journal of Animal Ecology, **73**, 599-611.

Kalinowski ST, Waples RS (2002) Relationship of effective to census size in fluctuating populations. Conservation Biology, **16**, 129-136.

Kekkonen J, Seppä P, Hanski IK, Jensen H, Väisänen RA, Brommer JE (2011) Low genetic differentiation in a sedentary bird: house sparrow population genetics in a contigous landscape. Heredity, **106**, 183-190.

Kimura M, Crow JF (1963) The measurement of effective population number. Evolution, **17**, 279-288.

Leberg P (2005) Genetic approaches for estimating the effective size of populations. Journal of Wildlife Management, **69**, 1385-1399.

Luikart G, Ryman N, Tallmon DA, Schwartz MK, Allendorf FW (2010) Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. Conservation Genetics, **11**, 355-373.

Nunney L (1993) The influence of mating system and overlapping generations on effective population size. Evolution, **47**, 1329-1341.

Nunney L (1995) Measuring the ratio of effective population size to adult numbers using genetic and ecological data. Evolution, **49**, 389-392.

Ovaskainen O, Hanski IA (2004) Metapopulation dynamics in highly fragmented landscapes.*in* I. A. Hanski and O. E. Gaggiotti, editors. Ecology, genetics, and evolution of metapopulations. Elsevier Academic Press, San Diego.

Palstra FP, Ruzzante DE (2011) Demographic and genetic factors shaping contemporary metapopulation effective size and its empirical estimation in salmonid fish. Heredity, **107**, 444-455.

Phillipsen IC, Funk WC, Hoffman EA, Monsen KJ, Blouin MS (2011) Comparative analyses of effective population size within and among species: Ranid frogs as a case study. Evolution, **65**, 2927-2945.

Pinheiro J, Bates D, DebRoy S, Sarkar D (2011) nlme: Linear and nonlinear mixed effects models. R Core team.

Pärn H, Jensen H, Ringsby TH, Sæther B-E (2009) Sex-specific fitness correlates of dispersal in a house sparrow metapopulation. Journal of Animal Ecology, **78**, 1216-1225.

Pärn H, Ringsby TH, Jensen H, Sæther B-E (2012) Spatial heterogeneity in the effects of climate and density-dependence on dispersal in a house sparrow metapopulation. Proceedings of the Royal Society Biological Sciences, **279**, 144-152.

R Development Core Team (2011) R: A languange and environment for statistical computing.*in* R. D. C. Team, editor. R Foundation for Statistical Computing.

Ringsby TH, Sæther B-E, Jensen H, Engen S (2006) Demographic characteristics of extinction in a small, insular population of house sparrows in northern Norway. Conservation Biology, **20**, 1761-1767.

Ringsby TH, Sæther B-E, Tufto J, Jensen H, Solberg EJ (2002) Asynchronous spatiotemporal demography of a house sparrow metapopulation in a correlated environment. Ecology, **83**, 561-569.

Schmeller DS, Merilä J (2007) Demographic and genetic estimates of effective population and breeding size in the amphibian Rana temporaria. Conservation Biology, **21**, 142-151.

Schwartz MK, Luikart G, Waples RS (2007) Genetic monitoring as a promising tool for conservation and management. Trends in Ecology & Evolution, **22**, 25-33.

Sæther B-E, Ringsby TH, Bakke Ø, Solberg EJ (1999) Spatial and temporal variation in demography of a house sparrow metapopulation. Journal of Animal Ecology, **68**, 628-637.

Tallmon DA, Koyuk A, Luikart G, Beaumont MA (2008) ONeSAMP: a program to estimate effective population size using approximate Bayesian computation. Molecular Ecology Resources, **8**, 299-301.

Templeton AR, Read B (1984) Factors eliminating inbreeding depression in a captive herd of speke's gazelle (*Gazella spekei*). Zoo Biology, **3**, 177-199.

Tufto J, Ringsby TH, Dhondt AA, Adriaensen F, Matthysen E (2005) A parametric model for estimation of dispersal patterns applied to five passerine spatially structured populations. American Naturalist, **165,** E13-E26.

Wang JL (2001) A pseudo-likelihood method for estimating effective population size from temporally spaced samples. Genetical Research, **78**, 243-257.

Wang JL (2005) Estimation of effective population sizes from data on genetic markers. Philosophical Transactions of the Royal Society B-Biological Sciences, **360**, 1395-1409.

Wang JL (2009) A new method for estimating effective population sizes from a single sample of multilocus genotypes. Molecular Ecology, **18**, 2148-2164.

Wang JL, Caballero A (1999) Developments in predicting the effective size of subdivided populations. Heredity, **82**, 212-226.

Wang JL, Whitlock MC (2003) Estimating effective population size and migration rates from genetic samples over space and time. Genetics, **163**, 429-446.

Waples RS (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. Genetics, **121**, 379-391.

Waples RS (2005) Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? Molecular Ecology, **14**, 3335-3352.

Waples RS (2006) A bias correction for estimates of effective population size based on linkage disequilibrium at unlinked gene loci. Conservation Genetics, **7**, 167-184.

Waples RS (2010) Spatial-temporal stratifications in natural populations and how they affect understanding and estimation of effective population size. Molecular Ecology Resources, **10**, 785-796.

Waples RS, Do C (2008) LDNE: a program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources, **8**, 753-756.

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary Ne using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. Evolutionary Applications, **3**, 244-262.

Waples RS, England PR (2011) Estimating contemporary effective population size on the basis of linkage disequilibrium. Genetics, **189**, 633-644.

Waples RS, Yokota M (2007) Temporal estimates of effective population size in species with overlapping generations. Genetics, **175**, 219-233.

Whitlock MC, Barton NH (1997) The effective size of a subdivided population. Genetics, **146**, 427-441.

Wright S (1931) Evolution in Mendelian populations. Genetics, **16**, 98-159.

Wright S (1938) Size of population and breeding structure in relation to evolution. Science, **87**, 430-431.

Data Accessibility

For microsatellite genotypes, genetic single-sample and temporal samples N_e-estimates and corresponding population characteristics see Dryad doi:10.5061/dryad.nb260. For individual-based data used to estimate demographic N_e see Dryad doi:10.5061/dryad.d02cn.

Author Contributions

HJ had the original idea and designed the study together with HTB. All authors performed the research. HTB analyzed the data with supervision and help from AMM, HJ, HP and IJH. AMM, BES, HJ, HTB and IJH wrote the paper with input from HP and THR.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Table S1: Overview of removed individuals/loci for the ONeSAMP estimation procedure.

Table S2: Overview of missing data/excluded data along with justification for removal of outliers.

Table S3: Estimates of effective population size using the single sample estimators LDNE and ONeSAMP.

Table S4: Estimates of effective population size using the temporal methods MLNE and CoNe.

Table S5: Estimates of effective population size for the total metapopulation for the two methods LDNE and MLNE.

Figure S1: The relationship between N_e and N_c or N_H within local populations for the four genetic estimators (LDNE, ONeSAMP, MLNE and CoNe) and the demographic estimator.

Appendix S1: Brief description of genotyping procedures.

Appendix S2: Description of model selection procedures.

Figure S2: The number of immigrants to each island population during three year intervals.

Figure S3: The relative increase in N_e-estimates from the MLNE(closed) method when the upper prior limit was increases from 2 times N_c to 40 times N_c, for temporal samples spaced either 1, 4 or 7 generations apart.

Figure Legends

Figure 1 The house sparrow metapopulation study system consisting of 18 island populations (shown in black) off the coast of Norway. The 15 islands included in this study are named.

Figure 2 Single sample effective population size estimates and population size over time in 15 house sparrow populations in Northern Norway. Census population size (N_c) is indicated as orange dots, N_e-estimates from LDNE are shown as green triangles, and N_e-estimates from ONeSAMP are shown as blue squares.

Figure 3 Temporal effective population size estimates and population size over time in 15 house sparrow populations in Northern Norway. Census population size (N_c) is indicated as orange dots, N_e-estimates from MLNE(closed) are shown in red, N_e-estimates from MLNE(open) are shown in green, and N_e-estimates from CoNe are shown in blue. For each estimator triangles show estimates based on samples separated by one generation, dots show estimates from samples separated by four generations, and squares show estimates from samples separated by seven generations.

Figure 4. Estimates of effective population size from (a) LDNE and (b) MLNE(closed) for the whole metapopulation, consisting of 15 insular house sparrow populations in Northern Norway. Census population size (N_c) is indicated as orange dots. For (a) LDNE the N_e-estimates for the total metapopulation ($metaN_e$) are shown in yellow whereas the sums of local population N_e (ΣN_e) are shown in green. For (b) MLNE(closed) the N_e-estimates for the total metapopulation ($metaN_e$) are shown in blue whereas the sum of local population N_e (ΣN_e) are shown in red; triangles show estimates based on samples separated by one

generation, dots show estimates from samples separated by four generations, and squares show estimates from samples separated by seven generations.

Tables

Table 1: Correlation between different estimators of effective population size ($\widehat{N}_e$) in a house sparrow metapopulation, and between $\widehat{N}_e$ and adult census population size (harmonic mean census population size N_H or annual census population size N_c). The upper left section gives the correlations between the temporal estimators (MLNE and CoNe) and demographic N_e; the lower right section gives the correlations for the single sample estimators (LDNE and ONeSAMP). P-values and the number of estimates (n) included in the correlation analyses are shown in brackets.

	$\widehat{N}_{e(MLNE,\ open)}$	$\widehat{N}_{e(MLNE,\ closed)}$	$\widehat{N}_{e(CoNe)}$	$\widehat{N}_{e(demographic)}$	$\widehat{N}_{e(LDNE)}$	$\widehat{N}_{e(ONeSAMP)}$
N_H	0.63 (p<<0.001, n=86)	0.83 (p<<0.001, n=86)	0.16 (p=0.269, n=51)	0.99 (p<<0.001, n=41)	-	-
$\widehat{N}_{e(MLNE,\ open)}$	-	0.69 (p<<0.001, n=86)	0.23 (p=0.122, n=51)	0.55 (p<0.001, n=41)	-	-
$\widehat{N}_{e(MLNE,\ closed)}$	-	-	0.36 (p=0.012, n=51)	0.83 (p<<0.001, n=41)	-	-
$\widehat{N}_{e(CoNe)}$	-	-	-	0. 40 (p=0.035, n=29)	-	-
N_c	-	-	-	-	0.63 (p<<0.001, n=65)	0.93 (p<<0.001, n=70)
$\widehat{N}_{e(LDNE)}$	-	-	-	-	-	0.62

	($p<<0.001$, $n=65$)

Table 2. Modeling variation in $\widehat{N}_e/N_c$ in a house sparrow metapopulation for the genetic estimators of effective population size (LDNE, ONeSAMP, MLNE and CoNe, respectively), as a function of population characteristics (sex ratio (*SR*), immigration rate (*m*), census population size (N_H or N_c), number of generations between samples (*g*; for the temporal estimators), population growth rate (*dN*/*dt*) and variance in population size ($\sigma^2_{N_c}$)). All models nested within the global models were tested (see Statistical analyses), however only a subset containing the highest ranked models are listed. *K* denotes the number of parameters, *L* is the log Likelihood of the model, AIC_C is Aikaike's information criterion for small sample sizes, Δ_i is the difference in AIC_C between the best model and model *i*, w_i is the Aikake weight of model *i*, and *ER* is the evidence ratio. The "best" model for each estimator is indicated in bold.

$\widehat{N}_e/N$ *and model*	*K*	*L*	AIC_C	Δ_i	w_i	*ER*
a) $\widehat{N}_{e(LDNE)}/N_c$						
$SR + m$	**2**	**-43.96**	**96.61**	**0.00**	**0.155**	**1.00**
$N_c + SR + m + N_c{*}SR$	4	-41.87	97.24	0.63	0.114	1.36
$SR + m + dN/dt$	3	-43.26	97.57	0.96	0.096	1.61
$N_c + SR + m$	3	-43.34	97.72	1.11	0.089	1.74
$N_c + SR + m + dN/dt + N_c{*}SR$	5	-40.99	98.02	1.41	0.077	2.01
b) $\widehat{N}_{e(ONeSAMP)}/N_c$						
N_c	**1**	**-20.41**	**47.19**	**0.00**	**0.227**	**1.00**
$N_c + \sigma^2_{N_c}$	2	-19.96	48.55	1.36	0.115	1.97
c) $\widehat{N}_{e(MLNE,\ closed)}/N_H$						
$N_H + SR + m + dN/dt$	**4**	**-93.70**	**200.47**	**0.00**	**0.072**	**1.00**
$N_H + SR + \sigma^2_{N_c} + m + dN/dt + N_H{*}m$	6	-91.35	200.56	0.09	0.069	1.04
$N_H + SR + \sigma^2_{N_c} + m + dN/dt$	5	-92.61	200.65	0.18	0.066	1.09
$N_H + SR + \sigma^2_{N_c} + m + dN/dt + N_H{*}SR + N_H{*}m$	7	-90.44	201.25	0.78	0.049	1.47
$N_H + SR + m + dN/dt + N_H{*}m$	5	-92.91	201.25	0.78	0.049	1.47
$N_H + SR + \sigma^2_{N_c} + m + N_H{*}m$	5	-92.93	201.30	0.83	0.048	1.50
$N_H + SR + \sigma^2_{N_c} + m$	4	-94.22	201.50	1.03	0.043	1.67
$N_H + SR + m + dN/dt + N_H{*}SR$	5	-93.23	201.89	1.42	0.035	2.06

$N_H + SR + m + dN/dt + N_H{*}SR + N_H{*}m$	6	-92.10	201.97	1.50	0.034	2.12
$N_H + SR + \sigma^2_{N_c} + m + dN/dt + N_H{*}SR$	6	-92.17	202.21	1.74	0.030	2.40
$N_H + SR + \sigma^2_{N_c} + m + N_H{*}SR + N_H{*}m$	6	-92.26	202.38	1.91	0.028	2.57
d) $\widehat{N}_{e(MLNE,\ open)}/N_H$						
$\boldsymbol{N_H + \sigma^2_{N_c}}$	**2**	**-132.80**	**274.03**	**0.00**	**0.200**	**1.00**
$N_H + SR + \sigma^2_{N_c}$	3	-132.54	275.80	1.77	0.082	2.44
$N_H + \sigma^2_{N_c} + m + N_H{*}m$	4	-131.41	275.85	1.82	0.080	2.50
$N_H + \sigma^2_{N_c} + m$	3	-132.60	275.92	1.89	0.077	2.60
e) $\widehat{N}_{e(CoNe)}/N_H$						
$\boldsymbol{N_H + m + dN/dt + N_H{*}m}$	**4**	**-70.64**	**155.44**	**0.00**	**0.232**	**1.00**
$N_H + \sigma^2_{N_c} + m + dN/dt + N_H{*}m$	5	-69.64	156.23	0.79	0.156	1.49

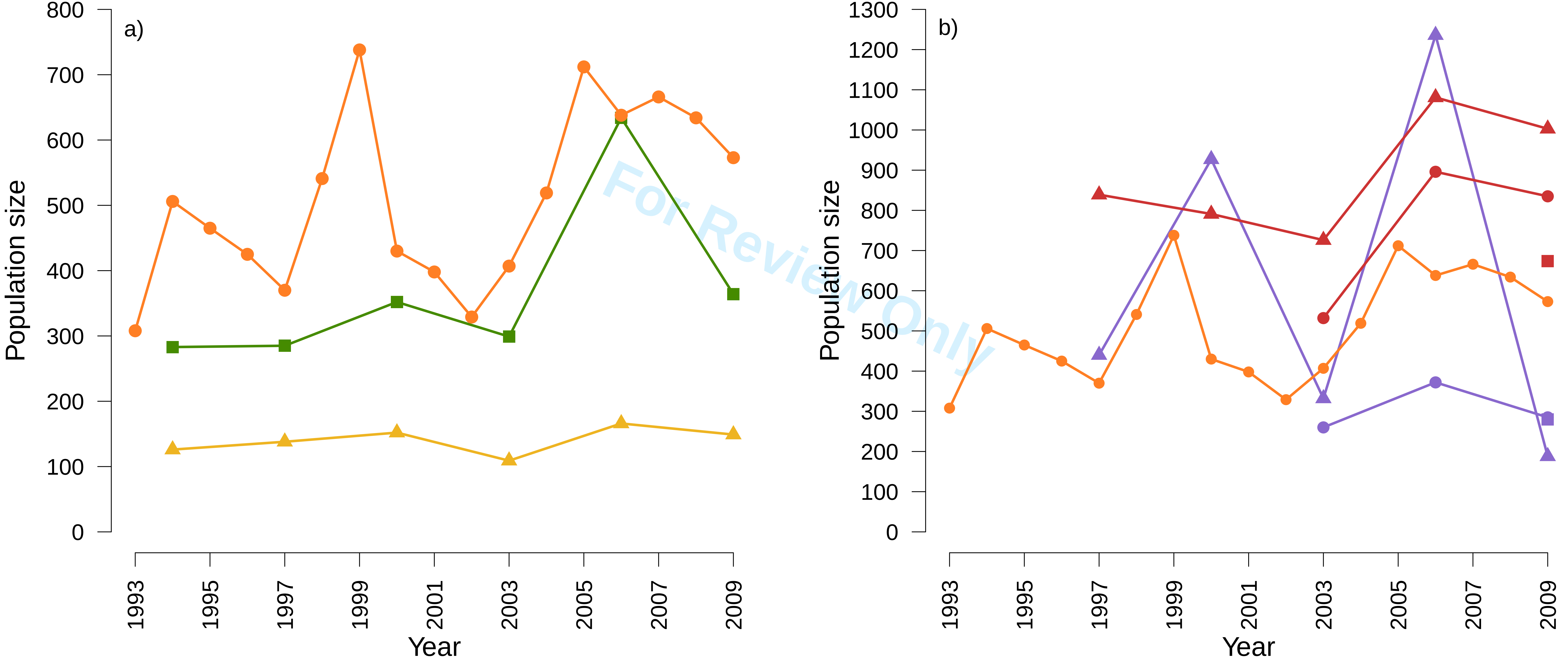
Table 3. Parameter estimates (coefficient) and standard error (*SE*) for the explanatory variables (sex ratio (*SR*), immigration rate (*m*), census population size (N_H or N_c), population growth rate (dN/dt) and variance in population size ($\sigma^2_{N_c}$)) in the "best" models explaining variance in a) $\widehat{N}_{e(LDNE)}/N_c$ and b) $\widehat{N}_{e(ONeSAMP)}/N_c$, c) $\widehat{N}_{e(MLNE,\ closed)}/N_H$, d) $\widehat{N}_{e(MLNE,\ open)}/N_H$ and e) $\widehat{N}_{e(CoNe)}/N_H$, respectively, in a house sparrow metapopulation.

Model parameter	β	*SE*	*p-value*
a) $\widehat{N}_{e(LDNE)}/N_c$			
Intercept	-0.502	0.368	0.178
SR	2.083	0.674	0.003
m	6.563	1.489	<0.001
b) $\widehat{N}_{e(ONeSAMP)}/N_c$			
Intercept	0.934	0.066	<<0.001
N_c	0.007	0.001	<0.001
c) $\widehat{N}_{e(MLNE,\ closed)}/N_H$			
Intercept	4.352	0.787	<0.001
N_H	-0.017	0.003	<0.001
SR	-2.992	1.369	0.032
m	7.004	2.646	0.010
dN/dt	0.260	0.120	0.033
d) $\widehat{N}_{e(MLNE,\ open)}/N_H$			
Intercept	3.170	0.200	<0.001
N_H	-0.031	0.004	<0.001
$\sigma^2_{N_c}$	0.002	0.000	<0.001
e) $\widehat{N}_{e(CoNe)}/N_H$			
Intercept	2.574	0.600	<0.001
N_H	-0.015	0.011	0.197
m	52.220	10.268	<<0.001
dN/dt	0.938	0.310	0.004
$N_H{*}m$	-0.549	0.248	0.032


Myken
Selvær
Nesøy
Gjerøy
Sundøy
Selsøyvik
Hestmannøy
Træna
Ytre Kvarøy
Indre Kvarøy
Onøy
Aldra
Lurøy
Lovund
Sleneset
0
5
10
20 km
0°
10°E
20°E
30°E
70°N
65°N
60°N
10°E
20°E


Aldra
Gjerøy
Hestmannøy
Indre Kvarøy
Lovund
Lurøy
Myken
Nesøy
Onøy
Selsøyvik
Selvær
Sleneset
Sundøy
Træna
Ytre Kvarøy
Population size
Year
1993
1995
1997
1999
2001
2003
2005
2007
2009


Aldra
Gjerøy
Hestmannøy
Indre Kvarøy
Lovund
Lurøy
Myken
Nesøy
Onøy
Selsøyvik
Selvær
Sleneset
Sundøy
Træna
Ytre Kvarøy
Population size
Year
1993
1995
1997
1999
2001
2003
2005
2007
2009
0
20
40
60
80
100
120
140
160
180
200
220
240
260
280
320
360
10
30
50
70
90


a)
b)
Population size
Year
0
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1993
1995
1997
1999
2001
2003
2005
2007
2009