

# Short-Term Genetic Changes: Evaluating Effective Population Size Estimates in a Comprehensively Described Brown Trout (*Salmo trutta*) Population

Dimitar Serbezov,<sup>\*,1</sup> Per Erik Jorde,<sup>\*,†</sup> Louis Bernatchez,<sup>\*</sup> Esben Moland Olsen,<sup>\*,†</sup> and L. Asbjørn Vøllestad<sup>\*</sup>

<sup>\*</sup>Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, Blindern, 0316 Oslo, Norway,

<sup>†</sup>Institute of Marine Research, Flødevigen, 4817 His, Norway, and <sup>‡</sup>Institute de Biologie Intégrative et des Systèmes, Université Laval, Québec QC G1V 0A6, Canada

**ABSTRACT** The effective population size ( $N_e$ ) is notoriously difficult to accurately estimate in wild populations as it is influenced by a number of parameters that are difficult to delineate in natural systems. The different methods that are used to estimate  $N_e$  are affected variously by different processes at the population level, such as the life-history characteristics of the organism, gene flow, and population substructure, as well as by the frequency patterns of genetic markers used and the sampling design. Here, we compare  $N_e$  estimates obtained by different genetic methods and from demographic data and elucidate how the estimates are affected by various factors in an exhaustively sampled and comprehensively described natural brown trout (*Salmo trutta*) system. In general, the methods yielded rather congruent estimates, and we ascribe that to the adequate genotyping and exhaustive sampling. Effects of violating the assumptions of the different methods were nevertheless apparent. In accordance with theoretical studies, skewed allele frequencies would underestimate temporal allele frequency changes and thereby upwardly bias  $N_e$  if not accounted for. Overlapping generations and iteroparity would also upwardly bias  $N_e$  when applied to temporal samples taken over short time spans. Gene flow from a genetically not very dissimilar source population decreases temporal allele frequency changes and thereby acts to increase estimates of  $N_e$ . Our study reiterates the importance of adequate sampling, quantification of life-history parameters and gene flow, and incorporating these data into the  $N_e$  estimation.

**T**HE effective population size ( $N_e$ ) is an essential concept in evolutionary and conservation biology as it determines the strength of stochastic evolutionary processes relative to deterministic forces (Crow and Kimura 1970). In the absence of gene flow, the rate of loss of genetic diversity via genetic drift is greater in populations with small  $N_e$ . The effective population size is, however, notoriously difficult to accurately estimate in wild populations as it is influenced by a number of parameters that are difficult to characterize in natural systems. A number of different methods have been developed for estimating  $N_e$ , and these are affected variously by different processes at the population level such as immigration, fluctuations in population size, population substructure, and life-history characteristics. When possible,

it is therefore important to compare  $N_e$  estimates obtained using different methods and elucidate how different processes affect the various methods for  $N_e$  estimation (Fraser *et al.* 2007).

Two main types of approaches have been used to estimate  $N_e$ : genetic and demographic. Genetic methods attempt to infer the magnitude of the effective population size by characterizing the genetic consequences of limited population size, whereas demographic methods depend upon measurement of demographic parameters that have theoretically been shown to influence  $N_e$ . The most common genetic approach is the so-called temporal method (Waples 1989). It is based on measurements of the amount of random genetic drift between generations that is inversely proportional to the effective population size (Nei and Tajima 1981). Although this method typically makes the simplifying assumption that generations are discrete, it is routinely applied also to species with overlapping generations. Overlapping generations are expected to bias  $N_e$  estimates because temporal fluctuations in allele frequencies follow more complicated

Copyright © 2012 by the Genetics Society of America  
doi: 10.1534/genetics.111.136580

Manuscript received November 11, 2011; accepted for publication March 22, 2012

<sup>1</sup>Corresponding author: Centre for Ecological and Evolutionary Synthesis, Department of Biology, P.O. Box 1066, University of Oslo, Blindern, 0316 Oslo, Norway. E-mail: d.r.serbezov@bio.uio.no

patterns when generations overlap (Jorde and Ryman 1995). Especially relevant is to explore short-term genetic fluctuations, as researchers are often interested in obtaining an estimate of  $N_e$  without having to sample a population interval several generations apart, which would be impractical, especially in long-lived organisms. With overlapping generations the amount of allele frequency change depends not only on the true effective size and the time between samples, but also on what population units are being sampled (e.g., single cohorts or aggregates of multiple cohorts) and the life-history parameters (age-specific birth and survival rates) linking these units to the effective size of the population under consideration (Jorde and Ryman 1995). Such complications underscore the need for empirical  $N_e$  studies founded on detailed knowledge of the organism's life history and age structure, to validate the various genetic estimation methods.

When generations overlap, an alternative to estimate the usual, per generation, effective size ( $N_e$ ) is to focus on a single breeding season and estimate the per-season effective number of breeders ( $N_b$ ) (*sensu* Waples 1990). In most real populations  $N_b$  will be less than the number of sexually mature individuals because of larger than binomial variance in family sizes. In semelparous organisms (*i.e.*, species that breed only once and die immediately after),  $N_e$  is simply the sum of  $N_b$  over a generation. It is, however, unclear how these two effective numbers relate in the more common iteroparous (repeat breeders) situation and it is thus useful to estimate both numbers for such species.

Most wild populations will be at least partially connected to other conspecific populations by migration. The resulting gene flow will contribute to the genetic constitution of the recipient population and affect the estimated  $N_e$ . Difficulties in measuring  $N_e$  also arise from the lack of adequately accounting for the spatial and metapopulation structure and clarity on how these affect  $N_e$ . Depending on the variation in size and productivity of subpopulations and the patterns of migration and extinction/recolonization rates, the sum of  $N_e$ 's of the subpopulations can either be greater or smaller than the effective size of the total (meta)population ( $N_{eT}$ ) (Nunney 1999; Palstra and Ruzzante 2011). With high levels of dispersal, the concept of a (sub)population  $N_e$  becomes increasingly meaningless since it is not acting as an independent unit (Nunney 2000). Fragmented populations may instead form dynamic nonequilibrium metapopulation systems with reduced total effective size ( $N_{eT}$ ) (Hedrick and Gilpin 1997). Quantifying all these effects is often difficult and our knowledge of the effective size in natural populations is consequently limited.

The sampling scheme itself can have pronounced effects on estimates of  $N_e$ . Sampling issues need to be considered carefully when applying a particular estimator for  $N_e$ , as different estimators make different assumptions on the structure of data and the conditions under which individuals are collected. For genetic methods, precision is restricted by sampling limited numbers of individuals and genes. Bias

may result from improper sampling of the population and unbiased estimates of allele frequencies require representative samples of individuals from each age class, which is rarely obtained from natural populations. In systems where migration occurs, it is also important that samples are drawn from the entire population and not only from subunits of the population (Nei and Tajima 1981; Waples 1989). Also, for the demographic methods representative offspring sampling is crucial for obtaining accurate estimates of family size variation. Accurate estimation of  $N_e$  thus requires careful consideration of a number of sources for potential errors.

In this study, we compare different  $N_e$  estimates in a semi-isolated resident brown trout population for which age-specific fertility and survival, as well as gene flow, have been estimated. The system has been thoroughly sampled from 2002 to 2006 and all sampled fish have been aged and genotyped. We could thus obtain allele frequency estimates for five annual (multicohort) samples and for eight consecutive single cohorts (1998–2006), and we used these to estimate genetic drift and effective size. In addition, we have a demographic  $N_e$  estimate (Serbezov *et al.* 2011) calculated from a parentage analysis study (Serbezov *et al.* 2010) for this population. Thus, we can compare the demographic  $N_e$  estimate with various genetic estimates presented herein. Specifically, we compare the demographic  $N_e$  estimate with genetic estimates obtained by using a method specifically developed for organisms with overlapping generations (Jorde and Ryman 1995), a method based on short-term genetic changes but also allowing for gene flow (Wang and Whitlock 2003), a method based on the relationship between randomly generated linkage disequilibrium and the effective population size (Hill 1981; Waples and Do 2008), and a maximum-likelihood coalescent-based method (Beaumont 2003). The comparison of several methods applied to data from this comprehensively described system allows us to assess how various population processes affect  $N_e$  estimates in natural populations.

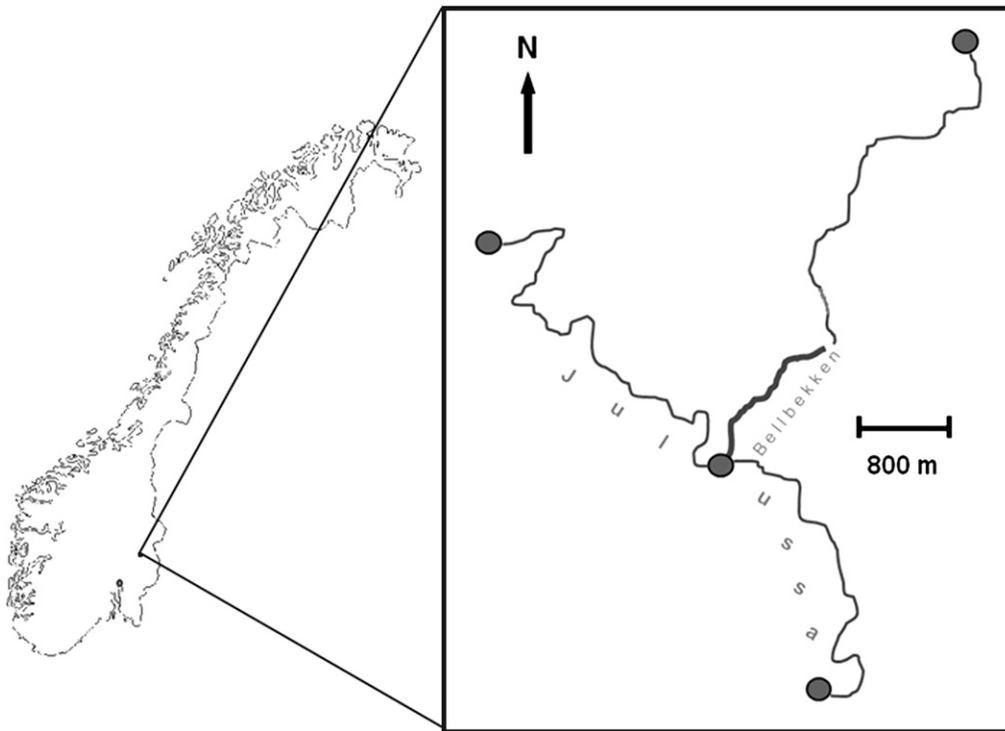
## Materials and Methods

### Brown trout

Brown trout (*Salmo trutta*) mate in fresh water and the eggs are commonly deposited in clean gravel (redds) in running water by the female. After hatching and depletion of yolk-sac reserves, the alevins emerge from the gravel and establish feeding territories, facing a period of high mortality risk (Elliott 1994). During later life stages, some trout remain in their natal stream whereas others migrate to larger systems—rivers, lakes, or the ocean to feed—and then home to their natal habitat to breed.

### Study site and data collection

Brown trout were sampled from a small forest stream, Bellbekken (Figure 1), in southeast Norway (N 61.15°, E 11.51°) during the period 2002–2006 (see Carlson *et al.* 2008 and Serbezov *et al.* 2010 for detailed information on



**Figure 1** Overview of the sampled locations. Bellbekken (1504 m) was exhaustively sampled in the period 2002–2007. Samples from a location 1450 m upstream of the focal area were collected in 2006 and from three locations in the larger river Julussa in 2005 and 2006 (shaded circles).

the stream and the sampling procedure). Briefly, a 1.5-km stretch of the stream was divided into 25 contiguous sections that were electrofished systematically and thoroughly during early summer (June) and autumn (late September to early October), starting in autumn 2002 and ending in autumn 2006. The downstream section starts at a small waterfall (section 1 below the waterfall, sections 2–25 above the waterfall), which should prevent upstream migration under most conditions, leading to weak but significant genetic differentiation between trout upstream and downstream the waterfall (Taugbøl 2008). Below the waterfall the stream enters the larger river Julussa. To map all potential sources of migration to Bellbekken we also sampled trout outside the focal 1.5-km study area (Figure 1). Specifically, during 2005 and 2006 brown trout were sampled from four additional sites: (1) at the confluence between Julussa and Bellbekken ( $n = 77$  trout), (2) 2.6 km upstream in Julussa ( $n = 51$ ), (3) 4.5 km downstream in Julussa ( $n = 61$ ), and (4) 1.5 km upstream in Bellbekken ( $n = 48$ ). Tissue samples were taken noninvasively from all fish, and individuals in Bellbekken were aged, sexed if mature, and marked as described previously (for details, see Serbezov *et al.* 2010).

#### Genetic analysis

All sampled fish (in total 4441 individuals) were analyzed with 15 microsatellite loci (details of DNA extraction and the PCR amplification can be found in Serbezov *et al.* 2010). The 15 loci were tested for null alleles with the Microchecker (Van Oosterhout *et al.* 2004) and CERVUS (Marshall *et al.* 1998) software packages, and no null alleles

were found. The year–site–loci samples did not show any systematic deviations from Hardy–Weinberg expectations and <10% of the samples were significant at the 5% level of confidence without allowance for multiple tests. Many of these were due to heterozygosity excesses, indicating that they are not caused by null alleles. Twelve percent of all individuals were genotyped more than once, allowing us to estimate the genotyping and scoring error rates (see Serbezov *et al.* 2010). We used Arlequin v3.1 (Excoffier *et al.* 2007) to quantify the spatial and temporal components of the genetic variance. This was done by regarding the consecutive (autumn) samples (from 2002 to 2006) as the temporal source of variation and the stream sections (2–25) as the spatial source of variation.

#### Population structure and migration

We used STRUCTURE 2.3.3 (Pritchard *et al.* 2000) to estimate the total number of populations in the pooled material from Julussa and Bellbekken. The Julussa sample was made up of all fish sampled there in the fall 2005 and 2006 (total of 213 individuals), and the Bellbekken sample consisted of the same number of individuals, randomly chosen from the Bellbekken fall 2005 sample. The optimal number of populations ( $K$ ) was determined from  $\Delta K$  likelihood evaluations (Evanno *et al.* 2005). As two distinct populations were inferred (see *Results*), we identified migrants into Bellbekken using GENECLASS2 (Piry *et al.* 2004). We conducted the assignment using the Bayesian probability method (Rannala and Mountain 1997) and Monte Carlo resampling for probability testing with 10,000 simulated individuals (type-I error,  $\alpha = 0.01$ ).

The genealogical history of alleles in Bellbekken and the larger Julussa was inferred considering two alternative models of population structure: a genetic drift model and a gene flow model. The drift model is that of an ancestral, panmictic population that has split into two isolated populations that diverge by random genetic drift alone. The gene flow model, on the other hand, assumes that the gene frequencies in the two populations are determined by a balance between genetic drift and gene flow (migration). Gene flow depends on the effective number of migrants,  $N_e m$  ( $N_e$  being the effective population size and  $m$  is the fraction of the population exchanged per generation), whereas the amount of drift depends on  $t/N_e$ , where  $t$  is the time since separation of the two populations. The *2mod* software (Ciofi *et al.* 1999) calculates the likelihood of each model on the basis of averaging over all possible values of  $N_e m$  and  $t/N_e$ .

Ongoing gene flow and migration rates were also estimated using BAYESASS v1.3 (Wilson and Rannala 2003). This Bayesian method uses the fact that immigrants show temporal linkage disequilibrium in their genotypes relative to members of the hosting population, thus allowing their identification as recent (first or second generation) immigrants.

### Demographic estimates of $N_e$

Demographic estimates of  $N_e$  were obtained using expressions derived by Hill (1972, 1979) and described in Serbezov *et al.* (2011). Briefly, we obtained estimates of variance in life-time reproductive success in Bellbekken by using an individual-based model using empirical estimates of sex- and age-specific reproductive success (Serbezov *et al.* 2010) and age-specific survival rates (Carlson *et al.* 2008). Age-specific maturation probabilities were also modeled on empirical data (Serbezov *et al.* 2011). By tracking each member of the cohort throughout its life, the model directly incorporates age structure in the mean and variance of life-time reproductive success and does so without assuming stable age distribution. This model was also run on a subset of the data after removing all individuals that were assigned as potential immigrants from Julussa.

### Temporal methods for estimating $N_e$

The effective population size ( $N_e$ ) of the brown trout population in the Bellbekken stream was also estimated from temporal variance in microsatellite allele frequencies. Four different data sets were used, *i.e.*, annual samples of all individuals caught in Bellbekken, with and without individuals assigned as putative migrants, and all individuals belonging to the same cohort, again with and without potential immigrants (Table 1). Because sampling was non-invasive and fish were marked and returned after sampling, a large number of individuals appear in more than one of the annual samples, whereas individuals were counted only once when constructing the cohort data sets (*cf.* the larger sums of annual samples than cohort samples in Table 1).

**Table 1** The number of sampled brown trout from Bellbekken at each fishing episode and the number of fish belonging to one of eight consecutive cohorts

	Sample size	Assigned as immigrants from Julussa	Sample size (excluding immigrants)
Fishing episode			
Fall 2002	741	105	636
Spring 2003	499	62	437
Fall 2003	591	56	535
Spring 2004	1051	79	972
Fall 2004	1018	97	921
Spring 2005	1017	75	942
Fall 2005	1103	98	1005
Spring 2006	631	41	590
Fall 2006	512	49	463
Cohort			
1998	166	16	150
1999	333	31	302
2000	340	66	274
2001	267	40	227
2002	740	49	691
2003	823	53	770
2004	703	49	654
2005	288	23	265

Exhaustive sampling was carried out twice each year (fall and spring) from 2002 to 2006, and scale readings or body length was used to classify fish to year of birth (*i.e.*, cohort). Potential immigrants from Julussa (Figure 1) were identified using GENECLASS v2 on the basis of statistical assignment to reference sample ( $n = 325$ ) from the Julussa river.

Separate analyses of samples with and without potential immigrants should counteract the consequences of the assumption of most methods of a single, isolated, population receiving no immigrants over the study interval. Most of the methods described below also assume, in addition to the population being closed to immigration, that generations are discrete and do not overlap. It is problematic to apply these methods to age-structured populations, such as the brown trout considered here, and for some methods it is not clear exactly what is being estimated from different groupings of the data (annual samples or cohorts). As annual samples are easier to obtain, most applications of these methods in the literature seem to be based on such “annual” samples and count the number of generations between samples (fractions of a generation in our case).

**The “standard” temporal method:** This method (*e.g.*, Waples 1989) calculates average (over loci) estimates of genetic drift ( $F$ ) and estimates  $N_e$  as

$$N_e = \frac{t}{2F}, \quad (1)$$

where  $t$  is the number of generations between the (two) samples. Here, we apply estimators  $F_k$  (Pollak 1983) and  $F_s$  (Jorde and Ryman 2007, Equation 9), separately for annual samples and consecutive cohorts, and correct estimates for the sampling scheme. Because individual fish were returned to the population just after sampling (fin clipping),

corrections should include population census size ( $N$ ) (Nei and Tajima 1981; Waples 1989). In the case of cohort analyses, the relevant census size is the number of juveniles (Jorde and Ryman 1995), which is typically much larger than either sample sizes or effective size. Hence, the factor  $1/N$  was ignored in the sample corrections of  $F_k$  and  $F_s$ , which was based on the respective formulas for sample “plan II” (Nei and Tajima 1981; Waples 1989). Statistical uncertainties in  $F$  and  $N_e$  were assessed by jackknifing over loci, using the leave-one-out procedure (Efron and Tibshirani 1993, Equation 11.3–11.4). The annual samples should yield an estimate of the annual effective size ( $N_a$ ), and it was converted to the effective size per generation ( $N_e$ ) by multiplying it by  $1/G$  (cf. Hill 1979, Equation 4), where  $G$  is the average generation length (see below).

**Maximum-likelihood coalescent-based method:** This method makes the same assumption of no mutations, no selection, no migration, and discrete generations, as the standard temporal method. Likelihood-based estimators are generally considered superior to the method-of-moments estimators (i.e., those relying on estimated amount of genetic drift,  $F$ ) in that they are apparently less sensitive to recent departures from migration–drift equilibrium (Tallmon *et al.* 2004). In addition, in populations with low  $N_e$  that do not follow the Wright–Fisher model (finite  $N$ , random mating and nonoverlapping generations), the gene genealogy might be better approximated by the coalescent (Berthier *et al.* 2002). We implemented this method with the TMVP software package developed by Beaumont (2003) for annual and cohort samples. It incorporates Bayesian prior information by making the user specify a maximum  $N_e$  (we used  $N_{e \text{ max}} = 1000$ ), resulting in a uniform prior distribution of  $N_e$  from 0 to  $N_{e \text{ max}}$ .

**Maximum-likelihood method [MLNE (Wang and Whitlock 2003)]:** Methods used to estimate  $N_e$  will be biased if the effects of immigration on allele frequencies are large relative to the effect of genetic drift. In addition to running all models with and without putative immigrants, we evaluated the impact of gene flow on  $N_e$  by applying Wang and Whitlock’s (2003) model incorporated in the MLNE software package, which jointly estimates  $N_e$  and the rate of gene flow ( $m$ ). This method is based on a pseudo-maximum-likelihood approach and uses temporal gene frequency changes in the focal population and the allele frequency of the potential source population to estimate effective size and migration rate separately. The method assumes that there is no selection and negligible mutation, that the source population is sampled, that allele frequencies in the source population and in the immigrants are stable, and that sampling does not affect the availability of reproductive individuals (Wang and Whitlock 2003).

**Accounting for age structure:** This was done following Jorde and Ryman (1995). Briefly, the method estimates  $N_e$  from observed temporal shifts in allele frequencies among consecutive cohorts (i.e., samples of individuals born 1 year apart). It does

this by accounting for the demographic characteristics of the population estimated from demographic parameters (the elements  $l_i$  and  $b_i$  of an ordinary Leslie population matrix) summarized into a correction factor  $C$  and generation length  $G$ ,

$$N_e = \frac{C}{G \times 2 \times F}, \quad (2)$$

where  $F$  is as in (1), but calculated from data on consecutive cohorts, and averaged over all consecutive cohort pairs (i.e., cohort pairs 1998 and 1999, 1999 and 2000, and so on).

To estimate the correction factors  $C$  and  $G$  we first used the field data to estimate annual survival ( $l_i$ ) and birth ( $b_i$ ) rates for each age class ( $i$ ). Age-specific survival rates were estimated from capture–mark–recapture data on brown trout in Bellbekken (Carlson *et al.* 2008). On the basis of estimated “winter” and “summer” season survival, age-specific monthly survival probabilities were estimated, and we used these estimates to calculate an average survival rate per year,  $S = 0.41$ . This estimate was subsequently used to calculate  $l_i = S^{i-1}$ , where  $i$  is the age class, counted from 1 ( $l_1 = 1$  per definition).

There was, however, some uncertainty in the age determination based on scales. In Bellbekken fish growth is slow and adjacent scale rings are often difficult to distinguish. A number of individuals that were not marked at first capture, or had lost their tags, were genotyped more than once and, as they had identical genotypes and compatible body length measurements, provided an estimate of the age-determination error rate. It was found that 36.5% of the fish that were aged  $\geq 3$  years were underestimated by 1 year and 19.0% were underestimated by 2 years. A corrected age distribution was obtained by iterating a computer simulation procedure in which 1 or 2 years are randomly added to the age of individual fish, in concurrence with the estimated proportions. Accordingly, age-specific birth rates ( $b_i$ ) were estimated in two steps. First, the probability of being mature at age  $i$ ,  $\text{prob}(\text{mature}_{(i)})$ , was estimated by fitting a logistic regression of the individual’s age on its maturity index (a binomial response variable). Second, the average number of offspring produced by parents of age  $i$  (age counted when their offspring were born, the following spring) was estimated on the basis of parentage assignment analysis from genetic data (Serbezov *et al.* 2010). We used a log-linear model to relate the age of the spawners to the number of assigned offspring. Because the slope ( $b$ ) of this regression was estimated for log-transformed data, the untransformed estimate is  $n_{\text{offspring}(i)} = i^b$ , where  $b$  is the (unweighted) mean of the estimated slopes for each sex. Finally, we estimated age-specific birth rates,  $b_i = \text{prob}(\text{mature}_{(i)}) \times \text{offspring}_{(i)}$ , which were normalized to yield a constant population size, by dividing by  $\sum l_i \times b_i$ . In both models, the logistic and log-linear, the uncertainty in the aging of the fish was accounted for by iterating the procedure described above 1000 times, and we used the median values (slope) from the simulated iterations.

**Table 2** Estimated survival ( $l_i$ ) and fecundity ( $b_i$ ) rates for each age class used to calculate the correction factor ( $C$ ) and generation interval ( $G$ ) for brown trout in Bellbekken

Age class	$l_i$	$b_i$
1	1	0
2	0.4100	0
3	0.1681	0.46
4	0.0689	2.12
5	0.0283	8.02
6	0.0116	21.15
7	0.0048	36.03
8	0.0019	46.50
9	0.0008	53.64

The resulting  $l_i$  and  $b_i$  values (Table 2) were used to estimate average generation length,  $G = \sum l_i \times b_i \times i$ , and correction factor  $C$  by iterating Equations 10–13 in Jorde and Ryman (1995) and substituting into their Equation 23 (using software “factorc”, available from <http://folk.uio.no/ejorde>). We arrived at estimates of  $C = 10.70$  and  $G = 5.73$  years. Finally, the effective size per generation,  $N_e$ , was estimated from allele frequency data from consecutive cohorts according to Equation 2, using data with and without putative immigrants.

#### **Linkage disequilibrium method for estimating $N_e$ [LDNE (Waples and Do 2008)]**

In addition to generating temporal change in allele frequencies (*i.e.*, random genetic drift), a limited effective size has other genetic consequences, including fluctuating associations of alleles at different loci (gametic phase or “linkage” disequilibrium). In an isolated population of constant size and stable structure, the amount of linkage disequilibrium (LD) can thus be used to measure  $N_e$  (Hill 1981). This method has advantages, such as requiring only a single sample of individuals. Like other methods it depends on a number of assumptions, in addition to those in common with the temporal methods above, also that loci are physically unlinked and independently segregating. Estimates of effective population size were obtained from annual samples and from single cohorts. Although to our knowledge this method has not been evaluated for age-structured species with overlapping generations, it is assumed that annual samples from the entire population should yield an estimate of per-generation  $N_e$ , whereas single-cohort data should characterize the effective number of breeders ( $N_b$ ) per spawning season (*cf.* Fraser *et al.* 2007). We therefore did not apply any corrections for overlapping generations to this method. The method was also applied to data excluding putative immigrants.

#### **The sibship assignment method for estimating $N_b$ [COLONY v2.0 (Wang 2009)]**

This is a hybrid approach with features of both the demographic and the genetic approaches. First, on the basis of a sample of individuals from a single cohort in the population, sibship assignments among cohort members are

obtained from multilocus genotype data. These assignments are then used to calculate half-sib and full-sib probabilities and then estimate  $N_e$ . This method assumes a population with discrete, nonoverlapping generations and, when it is applied to a sample of individuals taken at random from a single cohort in a population with overlapping generations, is supposed to give an estimate of the  $N_b$  that have produced this cohort. Wang (2009) suggested that this method for  $N_b$  estimation should be applicable to populations with either discrete or overlapping generations, random or nonrandom mating, and that it can be applied to estimate  $N_b$  in subpopulations with immigration (Wang 2009). The method is, however, fairly new and has to our knowledge not been extensively evaluated either in practical application or with computer simulations.

## **Results**

We successfully genotyped 3805 individual brown trout from Bellbekken belonging to eight consecutive cohorts (1998–2005), with at least 166 individuals from each (Table 1, Figure 2A). In addition, we genotyped 325 fish coming from different locations in the Julussa river (including station 1 from Bellbekken, which is below the waterfall and hence part of Julussa).

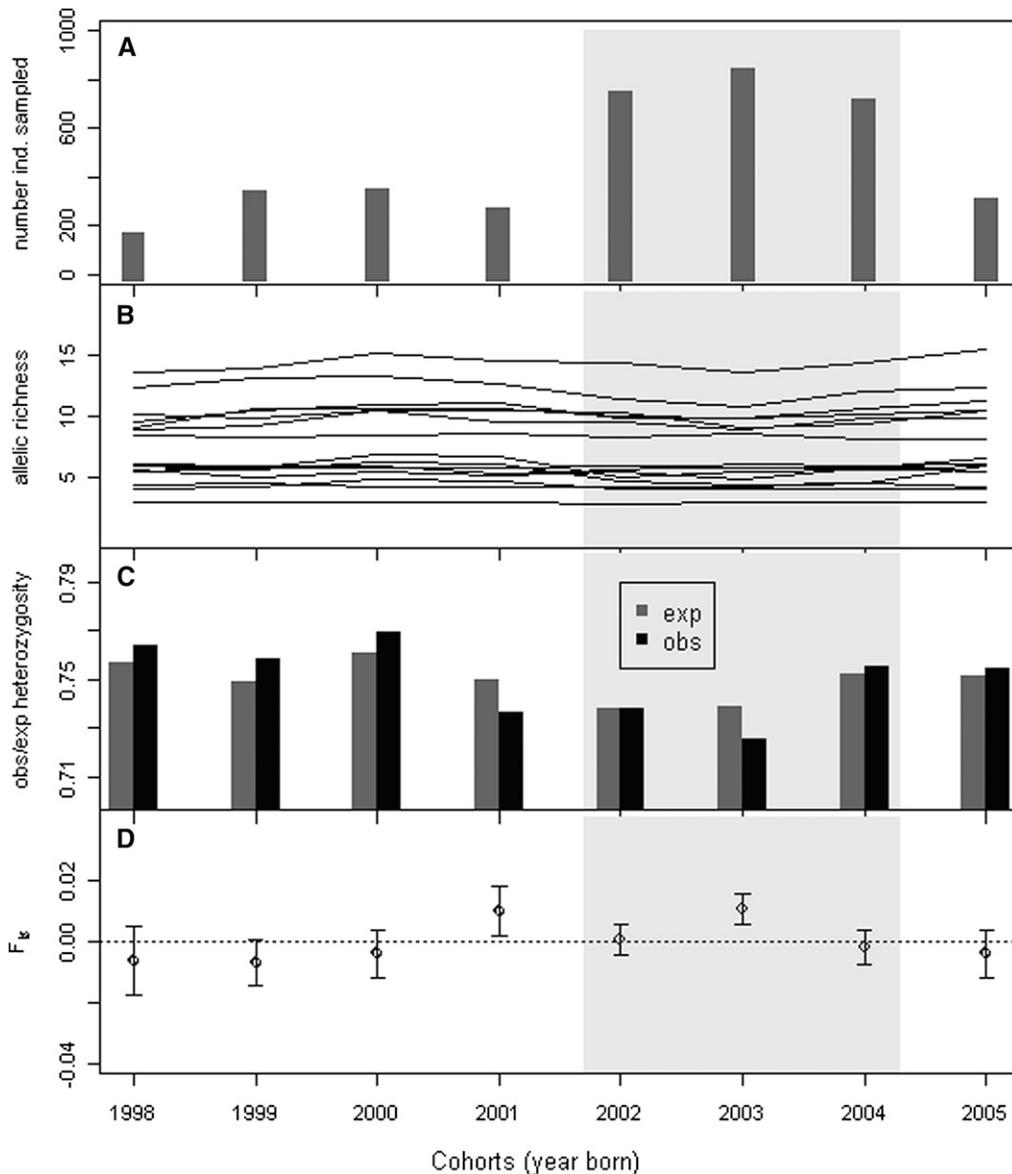
The number of alleles per locus over all samples ranged from 3 to 22 (mean = 11.9). The frequency of the most common allele ranged from 0.20 to 0.61 (mean = 0.35). Although the most common allele within a locus did change and there was some allele turnover, the allele frequency distribution was relatively stable across cohorts and there seemed to be no loss of rare alleles, leading to relatively stable temporal allelic richness values (El Mousadik and Petit 1996) (Figure 2C). The average observed heterozygosity in the eight consecutive cohorts was also very stable (0.73–0.76), but two of the eight cohorts had significantly positive  $F_{is}$  values (Figure 2D).

#### **Population structure and gene flow**

The temporal genetic variation among annual samples of brown trout in Bellbekken accounted for 0.02% and spatial differentiation accounted for 0.04% of the total genetic variation. An analysis of all pooled samples within Bellbekken showed a pattern of isolation by distance ( $\beta = 0.5$ ,  $r^2 = 0.25$ ,  $P < 0.05$ ) (L. A. Vøllestad, D. Serbezov, A. Bass, L. Bernatchez, E. M. Olsen, and A. Taugbøl, unpublished results).

Using STRUCTURE, the pooled material from Bellbekken and Julussa was best grouped into two populations ( $K = 2$ ). Bellbekken was moderately, but significantly, genetically differentiated from the larger Julussa ( $F_{st} = 0.018$ ,  $P < 0.05$ ). The population structure analysis overwhelmingly supported the gene flow model [ $p(\text{gene flow model}) = 1$ ].

Gene flow estimates between Bellbekken and Julussa are summarized in Table 3. The short-term gene flow estimates between Bellbekken and the “upstream” location obtained with BAYESASS were highly asymmetrical, with higher



**Figure 2** Characteristics for the eight consecutive cohorts of brown trout (*Salmo trutta*) sampled from Bellbekken. (A) The number of individuals sampled from each cohort. (B) Allelic richness of all 15 loci used in the analyses. (C) Expected and observed heterozygosity. (D) Inbreeding coefficient ( $F_{is}$ ). The zone with light shading marks the three cohorts for which we have estimated variance of individual reproductive success on the basis of parentage assignment analysis and used it for calculating a “demographic” estimate of  $N_e$ .

values in the Bellbekken upstream direction, compared to the other way around. Thus we chose, also considering the small sample size in the upstream location and the results from STRUCTURE, to treat Bellbekken and Julussa as a system of two populations that exchange genes, ignoring the sample from the upstream location. There was a high degree of congruence between short-term migration rate estimates between Bellbekken and Julussa obtained with BAYESASS and the MLNE estimate based on annual data.

#### Demographic method

The effective population size estimate ( $N_{ev}$ ) based on the variance in lifetime reproductive success was 104 [with a 95% confidence interval (C.I.) from 44 to 215]. The effective number of breeders per season obtained from empirical estimates of yearly reproductive success (Serbezov *et al.* 2012) ranged between 43 and 53 for the 3 consecutive years studied.

#### Temporal methods for $N_e$ estimation

**Estimates based on consecutive cohorts:** There was a consistent pattern of allele frequency change between cohorts (Figure 3), with highest  $F_s$  values observed at 3 years difference. For consecutive cohorts that were used to estimate effective size, the average  $F_s$  over loci was 0.01151, with a standard error (SE) of 0.00112. Using an average estimate of  $C/G = 1.87$  ( $C = 10.70$ ,  $G = 5.73$ ) in Equation 2, this corresponds to an estimated effective size per generation of  $N_e = 81$  (Table 4) for the Bellbekken brown trout population. There are two important sources of uncertainty in this point estimate. The first is the stochastic variation due to the limited sample sizes and number of screened loci. This uncertainty is quantified by the SE of  $F_s$  and a 95% C.I. for  $N_e$  from this effect is  $\sim C/(G \times 2 \times (F_s \pm 1.96 \times SE))$  or 68–101 (Table 4). Second, there is an uncertainty, and undoubtedly also natural variation, in the demographic parameters, leading to uncertainty in the calculated factor  $C/G$ . A sensitivity

**Table 3** Estimated ongoing migration rates ( $m$ ) of brown trout in and out of Bellbekken, based on Wang and Whitlock (2003) MLNE and Wilson and Rannala (2003) BAYESASS methods

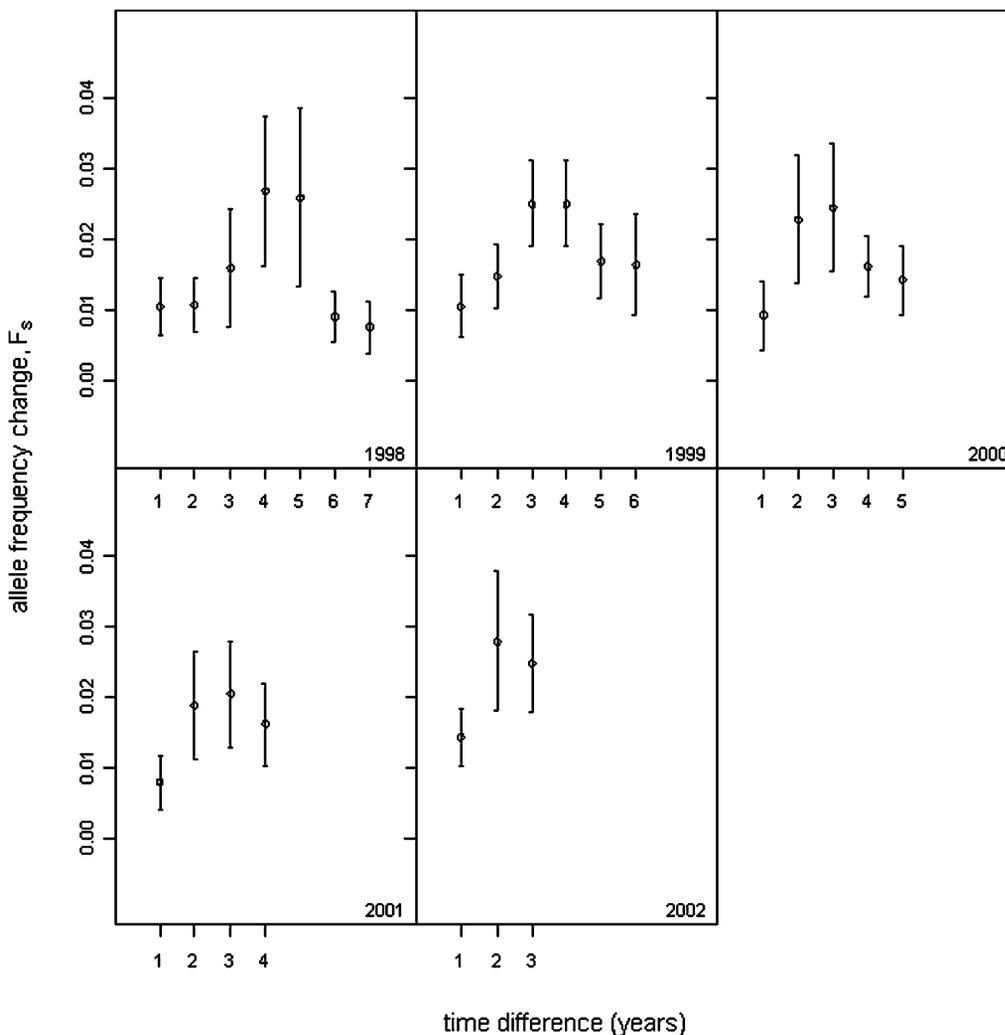
Method	Parameter estimated	Estimate (95% C.I.)
MLNE: annual samples	$m$ in Bellbekken	0.0083 (0.0082–0.0115)
MLNE: cohort samples	$m$ in Bellbekken	0.0746 (0.0614–0.0904)
BAYESASS	$m$ in Bellbekken from Julussa	0.012 (0.006–0.020)
BAYESASS	$m$ in Julussa from Bellbekken	0.019 (0.003–0.046)

analysis, varying demographic parameters (the annual survival rate and the slope of the logistic regression of number of offspring on parent age) around their point estimates, indicated that this ratio could vary between 1.52 and 2.00 over the likely parameter space (Figure 4). Taking this range as an indication of uncertainty in the estimated  $C/G$  ratio, the “confidence interval” for  $N_e$  would widen to  $\sim 55$ –107. This is a moderate adjustment to the original confidence interval and we have not included it in the results tabulated in Table 4.

Compared to the  $F_s$ -based method, the  $F_k$ - and the ML-based methods yielded somewhat higher estimates of  $N_e$

when based on consecutive cohort data and adjusted with the factor  $C/G$  (cf. Table 4). These ranged from the standard temporal method (Pollak 1983) estimate of 105 to 326 for the ML method of Wang and Whitlock (2003), with an intermediate estimate of 141 for the coalescent method of Beaumont (2003).

**Estimates based on annual samples:** Annual (autumn) samples of mixed-age composition contain partially overlapping sets of individuals, as many individuals appear in  $>1$  year and consequently fluctuated considerably less in allele frequencies over time than did cohort samples. The



**Figure 3** Allele frequency differences [ $F_s$ , adjusted for sampling (Jorde and Ryman 2007, Equation 13)] between pairs of cohorts (born 1998–2002) of brown trout in Bellbekken. The year in each plot refers to birth (hatching) year of the cohort that is being compared to subsequent cohorts (e.g., the 1998 cohort is compared to cohorts born in 1999, 2000, and up to 2005, seven comparisons altogether). Vertical bars refer to 95% confidence intervals for the point estimated, calculated by jackknifing over loci.

**Table 4** Estimated per-generation and seasonal effective sizes of brown trout in Bellbekken by different methods and data

Method	Correction <sup>a</sup>	All individuals		Excluding immigrants	
		$N_x$	95% C.I.	$N_x$	95% C.I.
Estimates of per-generation $N_e$					
Data: consecutive annual (autumn) samples					
Moments, $F_k$ (Pollak 1983)	1/G	121	89–176	106	77–153
Moments, $F_s$ (Jorde and Ryman 2007)	1/G	81	58–143	69	47–128
ML, closed (Wang and Whitlock 2003)	1/G	127	106–154	NA	
ML, open (Wang and Whitlock 2003)	1/G	106	89–130	NA	
Coalescent (Beaumont 2003)	1/G	87	66–111	91	69–117
Data: single annual (autumn) samples					
LD (Waples and Do 2008)	None	100	58–142	93	53–132
Data: consecutive cohort samples					
Moments, $F_k$ (Pollak 1983)	C/G	105	93–117	103	90–115
Moments, $F_s$ (Jorde and Ryman 2007)	C/G	81	68–101	78	65–98
ML, closed (Wang and Whitlock 2003)	C/G	326	305–351	NA	
ML, open (Wang and Whitlock 2003)	C/G	221	206–237	NA	
Coalescent (Beaumont 2003)	C/G	141	128–144	143	135–144
Data: demographic (Serbezov <i>et al.</i> 2010)					
Lifetime $V_k$ (Hill 1979)	None	104	44–215	96	45–205
Estimates of effective number of breeders $N_b$					
Data: single cohort samples					
LD (Waples and Do 2008)	None	53	22–84	52	17–87
Sibship (Wang and Whitlock 2003)	None	56	15–97	NA	
Data: demographic (Serbezov <i>et al.</i> 2010)					
Seasonal $V_k$	None	40	32–49	31	27–35

<sup>a</sup> Consecutive annual samples yield an estimate of drift per year. The tabulated estimates of per-generation effective size ( $N_e$ ) were calculated by multiplying the annual estimate by 1/G (Hill 1979, Equation 4), where  $G = 5.73$  is the average generation length. For consecutive cohorts the estimate refers to temporal allele frequency shifts among cohorts, and  $N_e$  estimates have been corrected by multiplying by the factor C/G or 1.87 (Equation 2). Estimates of per-season effective number of breeders ( $N_b$ : LD, sibship, and demographic methods only) were not adjusted. C.I. is the 95% confidence interval for the point estimates, either as given directly by the respective software or as calculated from standard errors for the relevant statistic ( $F$  or  $1/2N_e$ ), assuming normal distributed statistics, and was corrected with 1/G or C/G as for the corresponding point estimate.

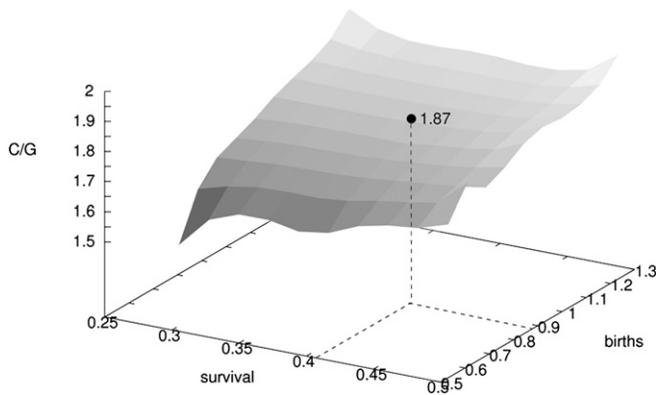
amount of allele frequency change from year to year ( $F_s$ , calculated from all consecutive pairs of years) averages 0.0011 (SE = 0.00024) or only approximately 1/10th the magnitude of cohort fluctuations (0.0115: see above). We used only the autumn samples when estimating annual drift and  $N_e$ , as also including spring samples would complicate the analyses of drift without contributing much in the way of new data as most individuals that were sampled in the spring were also found in the autumn samples. On the basis of the estimated annual drift among consecutive annual autumn samples, the two moments-based methods yielded estimates of  $N_e$  [adjusted for generation length by multiplying with 1/G (Hill 1979, Equation 4)] of 121 and 81, both very similar to the corresponding estimates based on cohort data (105 and 81, respectively; Table 4). Indeed, all methods were found to yield fairly similar results for annual data, ranging from 87 for the coalescent method to 127 for the ML method. These all compare well with the previously demographic estimate of  $N_e = 104$  (95% C.I. = 44–215) for this brown trout population. In the case of annual samples, however, the confidence intervals for the point estimates are typically broader than those based on cohort data.

**Estimates of effective number of breeders,  $N_b$ :** Estimates obtained using linkage disequilibrium data based on cohorts

(*i.e.*, single-year classes) were on average about half the magnitude of the corresponding estimate of  $N_e$  based on annual data (mean  $\pm$  95% C.I. =  $56.1 \pm 18.8$  vs.  $101.0 \pm 20.4$ ,  $t$ -test = 4.7, d.f. = 15,  $P = 0.0003$ ; Table 4; Figure 5) The only other method that gave an estimate of  $N_b$ , the sibship method, gave a very similar point estimate of 53. Both genetic estimates are thus in fair agreement with, but slightly higher than, the previous demographic estimate of the effective number of breeders per season of 40 (Table 4) for this population.

#### **The effect of gene flow on estimated $N_e$**

The procedures for estimating  $N_e$  and  $N_b$  were repeated on reduced data sets (annual samples or cohorts) with individuals identified by GENECLASS v2.0 as potential migrants removed (6.8% of all individuals were removed). There was no difference in the estimated proportion of immigrants between the sexes (test for equality of proportions:  $\chi^2 = 1.6$ , d.f. = 1,  $P = 0.2$ ). Removing putative migrants resulted in somewhat lower effective population size estimates for all methods (2–16% for different methods), except for the coalescent method that yielded very similar estimates with and without immigrants (*cf.* Table 4). Randomly removing the same number of individuals as the number of putative immigrants did not change the  $N_e$  estimates obtained using these



**Figure 4** Sensitivity analysis of the correction factor  $C/G$  over a range of demographic parameter values, for brown trout in Bellbekken. “Survival” is the annual survival rate (assumed to be constant for each age) and “births” refers to the slope of the logistic regression of number of offspring on parent age. The range of demographic values used in the analysis is a generous interpretation of the uncertainties of the estimated values given in Table 2. The solid circle at  $C/G = 1.87$  is the value that was used when estimating  $N_e$  from temporal allele frequency shifts among consecutive cohorts.

methods (data not shown), indicating that the slight reduction seen in the other methods is an effect, although modest, of immigration and not of reduced sample sizes.

## Discussion

In this study we have investigated short-term genetic changes in an exhaustively sampled and demographically mapped stream-resident brown trout population. Despite relatively small population size, we documented high levels of genetic diversity and no evidence of inbreeding. We then estimated the short-term effective population size ( $N_e$ ) using different genetic methods, all based on a number of assumptions, and compared the results with an estimate obtained using demographic data (Serbezov *et al.* 2011). There was broad congruence among estimates based on different genetic methods, as well as between genetic and demographic estimates. Some effects of violating the assumptions of the different methods were nevertheless apparent, as discussed below.

### Genetic variation and genotypic frequencies

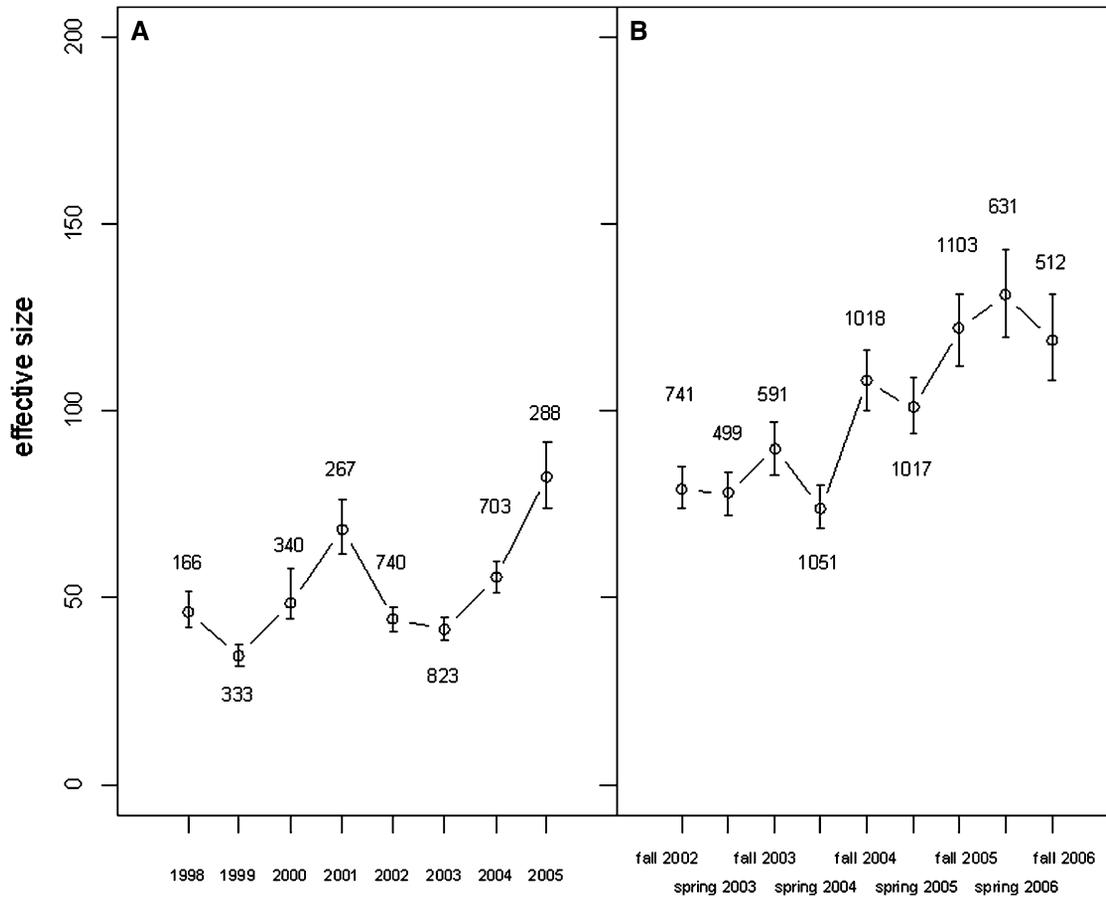
Low-frequency alleles were the most common allele frequency class, indicating that the population had not undergone any recent severe bottleneck. This was also confirmed by the fact that bottleneck tests were negative (results not shown). Also, the allele frequency distribution did not change significantly during the eight consecutive brood years studied, indicating a relatively stable number of spawners contributing each season. There did not seem to be any reduction of allele richness or level of heterozygosity, indicating that there were no inbreeding processes in our system during the period studied (8 consecutive brood years), most likely as a result of ongoing gene flow from

the larger river Julussa. That gene flow is a major factor preserving genetic variability is suggested by the fact that most of the populations of brown trout that have been reported to display a low degree of genetic variation are from systems where immigration is restricted; *i.e.*, the population is either physically isolated or located in the uppermost part of the drainage (Marshall *et al.* 1992; Prodohl *et al.* 1997; Carlsson and Nilsson 2001). In addition, the apparent demographic stability and the flexible mating system that involves both polygamy and iteroparity (Serbezov *et al.* 2010) would further help preserve the number of alleles and the high levels of heterozygosity.

### Distribution of genetic variation in time and space

The temporal genetic variation of brown trout in Bellbekken accounted for 0.02% of the genetic variability. This result is comparable to a previous study of brown trout, also based on a short time span [0.03% (Ryman 1983)]. The degree of spatial differentiation accounted for 0.04% of the total variation. This is approximately the same value as the one estimated in another study of brown trout within pairs of interconnected lakes with apparently high levels of gene flow (Ryman *et al.* 1979). Nevertheless, there was no clear indication that our stream-resident population is divided into lasting subpopulations. Analysis of the 2002–2004 cohorts revealed that although there can be some significant  $F_{st}$  values between the 24 sections of the stream, this differentiation is not consistent between years (data not shown). Also, there was no consistent significant heterozygote deficiency relative to the Hardy–Weinberg genotypic proportions at any locus, and indeed there were a number of cases of heterozygote excess, the opposite of what is expected from population subdivision. The occasional significant  $F_{st}$  values are apparently a result of a large proportion of the sample being from a stream section that consists of offspring belonging to the same family before they have dispersed. Brown trout offspring from a family do tend to cluster together in Bellbekken until their second summer (L. A. Vøllestad, D. Serbezov, A. Bass, L. Bernatchez, E. M. Olsen, and A. Taugbøl, unpublished results), although offspring belonging to the same families could be found all over the stream (Serbezov *et al.* 2010). Because of the lack of a permanent substructure, we treat our target as a single population. However, as discussed by Nei and Tajima (1981), the effect on the  $N_e$  estimates of sampling more than one genetically distinct population depends on the sampling strategy: if all populations are proportionately represented, the estimated effective size will be that of the total (combined) population. Our intensive and spatially uniformly distributed sampling should make sure that this is achieved to a satisfactory degree.

The magnitude of the  $F_s$  values, the amounts of allele frequency change obtained from annual samples, increased rather constantly between years (data not shown). This pattern indicates that we are indeed monitoring genetic drift in this study, as potentially confounding processes such as



**Figure 5** Estimates of effective population size ( $N_e$ ) of brown trout in Bellbekken based on linkage disequilibrium (LD) data. Shown are (A) estimates (with 95% C.I.) based on year cohorts were born in eight successive spawning seasons (1998–2006), and (B) estimates based on all individuals caught in nine consecutive fishing episodes (fall 2002 to fall 2006). The numbers next to the plotted points are the number of individuals ( $n$ ) the estimates are based on.

fluctuating selection, migration, and nonrandom sampling [e.g., kin sampling (Allendorf and Phelps 1981; Hansen *et al.* 1997)] are not expected to result in a linear increase with time. Also, in the cohort data, even though it hardly covers one brown trout generation, a consistent pattern of low  $F_s$  values was observed at about one generation (5.73 years) and a high value at 3 years (about half a generation) (*cf.* Figure 3). These patterns are in agreement with the theoretical expectations for selectively neutral alleles (*cf.* Jorde 2012) and observations lend support to the presumption that the main cause of the observed allele frequency changes is limited effective population size, *i.e.*, genetic drift.

#### Assumptions of genetic methods of estimating $N_e$

Genetic methods have been employed to circumvent the need to estimate  $N_e$  from demographic data. All such methods, however, make a variety of assumptions that are almost always broken to some degree. All assume no mutation or selection on the markers used in the analyses. As our study covers a very short period of time, hardly covering a single generation, we can safely ignore the effect of mutation. The

markers used in this study also do not seem to be under selection. In general, however, these methods also include several of the following assumptions: (i) adequate sample sizes or number of loci utilized, (ii) knowledge of the extent of gene flow as well as knowledge of source populations contributing migrants, (iii) no population size fluctuations over the period  $N_e$  was estimated, (iv) discrete generations, and (v) sufficient number of generations separating temporal samples. Violation of these assumptions could lead to a bias in the estimates. Nevertheless, the approximate congruence between most estimates suggests that we have arrived close the “true”  $N_e$  value using these methods. Slight exceptions to the general congruence are the estimates obtained by the ML method of Wang and Whitlock (2003) and to a lesser extent the coalescent method of Beaumont (2003), when using cohort sample data. Possibly, the adjustment for overlapping generations, by multiplication with  $C/G$ , is not appropriate for these methods. This issue needs further study.

Despite the general congruence among most methods and data sets, some effects of violating underlying assumptions are apparent. For example, assuming a closed

population when the focal population is not completely isolated is seen to affect systematically all methods and tends to bias estimates of  $N_e$  upward (except for the coalescent approach, Table 4). Removing putative immigrants or applying the only method specifically allowing for immigrants (the ML method) led to a reduced estimate of  $N_e$ , by a modest 2–16%. Bias introduced by uncritical adoption of the discrete generation model to age-structured species with overlapping generations could also bias estimates (Waples and Yokota 2007; Palstra and Ruzzante 2008; Waples *et al.* 2011). In this case bias depends on what population segments are being sampled, and we investigated two scenarios: sampling of single, consecutive cohorts and annual, multicohort samples. In either case similar, and presumed unbiased, estimates can be obtained (Table 4) if measurements are appropriately adjusted (employing correction factors  $C/G$  and  $1/G$ , respectively). Failure to adjust estimates would lead to corresponding bias. Bias with cohort samples arises mainly because the signals of genetic drift will strongly be affected by the pattern of allele frequency covariance among cohorts generated by age structure (Jorde and Ryman 1995; Jorde 2012). This pattern depends on the population's survival ( $l_i$ ) and fecundity ( $b_i$ ) values (Table 2), but the resulting  $C/G$  surface is quite flat over a wide range of parameter values (*cf.* Figure 4) and the  $N_e$  estimates are consequently little affected by the exact magnitude of these life-history values. Hence, studies lacking adequate data for estimating local demographic parameters may nevertheless use this method to correct bias by adopting published values from conspecific populations (*cf.* Jorde and Ryman 1996; Heggenes *et al.* 2009).

The most commonly used measures of allele frequency change,  $F_c$  (Nei and Tajima 1981) and  $F_k$  (Pollak 1983), tend to be downward biased (thus biasing upward the  $N_e$  estimates), when allele frequencies are close to zero or unity (Waples 1989). This bias was systematically characterized by Jorde and Ryman (2007) and an alternative estimator was proposed ( $F_s$ ). Although there is a trade-off between accuracy and precision when selecting estimators for genetic drift and effective size for the temporal method, computer simulations have shown that the  $F_s$  estimator yields unbiased estimates compared to those based on  $F_c$  and  $F_k$  estimators (Jorde and Ryman 2007). In our results, the  $F_k$  (and  $F_c$ ; data not shown)-based estimates gave indeed higher values than those based on the unbiased  $F_s$  (Table 4). This is probably so because allele frequencies in our study are not uniformly distributed, as would usually be the case in data from natural systems. Hence, the  $N_e$  value obtained by the  $F_s$  temporal method is probably closer to the true  $N_e$  value. On the other hand, the cohort model for iteroparous species may also underestimate  $N_e$  if there is a correlation between reproduction and future survival (Jorde and Ryman 1995). This is apparently because the variance in lifetime reproductive success is reduced (and true  $N_e$  therefore increases) as fewer breeders can reproduce more than once. Simulation studies show that this is indeed the case for our system

(Serbezov *et al.* 2011), and this might explain why the moments-based estimates from cohort data are somewhat lower than the rest. Also, genetic estimates of  $N_e$  derived from models for overlapping generations are sensitive to the confidence with which individuals have been assigned to their respective cohorts, on the basis of age determination from fish scales. Random errors in the assignment will likely introduce an upward bias into  $N_e$ , since a mixture of individuals of different ages in the presumed cohort will dilute the signals of genetic differentiation actually present among the cohorts. This might be the case in our study for older individuals (Serbezov *et al.* 2011).

The precision of the  $N_e$  estimates depends on the strength of the drift signal relative to sample noise. In species with overlapping generations these signals would be stronger between consecutive cohorts, compared to (mixed-age) samples between consecutive years. Consequently, the estimates obtained using cohort data typically had narrower confidence limits compared to the estimates obtained using annual samples (*cf.* Table 4).

We obtained  $N_e$  estimates from the LD method that were similar to those of the other methods and to the estimate based on demographic data (*cf.* Table 4). Recent theoretical and simulation results indicate that immigrants are expected to increase estimated  $N_e$  from this method, but only slightly so (Waples and England 2011). This prediction was confirmed in our empirical estimates that, as in most other methods, were a little lower with immigrants removed from the data (*cf.* Table 4). Moreover, the nine consecutive temporal samples gave rather congruent estimates: the  $N_e$  values all fell within the rather narrow range from 74 to 131. Our results thus suggest that the LD method of estimating  $N_e$  gives plausible estimates, at least when based on sufficiently large samples, and the validity of the estimates is confirmed by replicate temporal sampling. In the case of sufficient sampling, temporal replication of the LD method can be used to evaluate potential changes in  $N_e$  for genetic monitoring, and the effective size seemed to be quite stable over the study interval for the Bellbekken population (Figure 5).

### **Gene flow and its effect on estimated $N_e$**

In this study, we have analyzed genetic changes in a system where all potential gene flow source populations have been mapped and sampled. There was congruence in  $m$  estimates obtained using the Wilson and Rannala (2003) estimate and the Wang and Whitlock (2003) ML estimate based on annual sample data ( $m \approx 0.01$ , Table 3). The ML estimate based on cohort sample data was higher ( $m \approx 0.07$ ) and, as discussed in connection with the  $N_e$  estimates, there may be problems with applying the ML method to such cohort data. Computer simulations will be useful for exploring this issue in more detail.

Current analytical limitations complicate investigating the effects of gene flow on  $N_e$  over short time spans. Nevertheless, by removing potential immigrants from the  $N_e$  calculation with the temporal methods, the LD method (Waples and Do 2008), although not the coalescent method

(Beaumont 2003), resulted in slightly lower  $N_e$  estimates than the ones obtained on the basis of allele frequencies from all individuals sampled in Bellbekken (cf. Table 4). Wang and Whitlock (2003) have argued that, over few generations, migration increases allele frequency changes relative to change due to drift alone but, over many generations, migration has the opposite effect of stabilizing allele frequencies. Fraser *et al.* (2007) argued that, even in the short term, migration from a genetically similar population would not necessarily increase allele frequency change relative to drift alone. In the present situation, we found that including immigrants in the data set led to a slight decrease in the allele frequency change and a corresponding increase in estimated  $N_e$  (cf. Table 4). Similarly, the  $N_{e-CLOSED}$  (no migration) estimates were correspondingly higher than the  $N_{e-OPEN}$  (with migration) ones from the ML method (Wang and Whitlock 2003). Although this method assumes constant migration from an infinite source of fixed allele frequency, which is hardly the case in any natural systems, including ours, this does not seem to affect much our  $N_e$  or  $m$  estimates, possibly because migration was low. Environmental instability, fluctuations in demography, and life-history variation are likely factors influencing patterns of gene flow in a system. To fully understand the impact of gene flow on evolutionary dynamics of a system it is important to consider how gene flow varies over short and longer time spans (Hare *et al.* 2011).

### Conclusions

Many genetic methods have been proposed to estimate  $N_e$  (and  $N_b$ ) in natural populations and these make somewhat different assumptions regarding population and sample characteristics. On the basis of a comprehensive sampling effort in a natural brown trout population system, our findings demonstrate general congruence among estimates obtained by the various genetic methods and from demographic data. This congruence raises confidence in the general usefulness of genetic estimates of effective population size, also when the underlying assumptions are not fully met, as is almost always the situation with natural populations. The strength of the present study is in the extensive sampling effort, yielding relatively narrow confidence intervals for point estimates and thus allowing meaningful comparison among methods. Obviously, the true  $N_e$  is not known in this or other natural populations, and computer simulations with known  $N_e$  will be useful for evaluating methods over a wider range of parameter values. The present empirical study adds realism to such evaluations and provides a useful point of departure for future computer simulations.

A data file with the microsatellite genotypes, as well as the age and the place individuals were sampled, used in the analyses in this article is freely available at <http://dx.doi.org/10.5061/dryad.44pm4pq6>.

### Acknowledgments

We thank the large number of graduate students that have participated in fieldwork over the years. The large amount

of laboratory work would have been impossible without the kind assistance of Lucie Papillon, Vicky Albert, and Annette Taugbøl. We thank an anonymous reviewer for valuable comments. We thank the Norwegian Research Council for financial support.

### Literature Cited

- Allendorf, F. W., and S. R. Phelps, 1981 Use of allelic frequencies to describe population structure. *Can. J. Fish. Aquat. Sci.* 38: 1507–1514.
- Beaumont, M. A., 2003 Estimation of population growth or decline in genetically monitored populations. *Genetics* 164: 1139–1160.
- Berthier, P., M. A. Beaumont, J. M. Cornuet, and G. Luikart, 2002 Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: A genealogical approach. *Genetics* 160: 741–751.
- Carlson, S. M., E. M. Olsen, and L. A. Vøllestad, 2008 Seasonal mortality and the effect of body size: A review and an empirical test using individual data on brown trout. *Funct. Ecol.* 22: 663–673.
- Carlsson, J., and J. Nilsson, 2001 Effects of geomorphological structures on genetic differentiation among brown trout populations in a Northern Boreal river drainage. *Trans. Am. Fish. Soc.* 130: 36–45.
- Ciofi, C., M. A. Beaumont, I. R. Swingland, and M. W. Bruford, 1999 Genetic divergence and units for conservation in the Komodo dragon *Varanus komodoensis*. *Proc. R. Soc. Biol. Sci. Ser. B* 266: 2268–2274.
- Crow, J., and M. Kimura, 1970 *An Introduction to Population Genetics Theory*. Harper & Row, New York.
- Efron, B., and R. J. Tibshirani, 1993 *An Introduction to the Bootstrap*, pp. XVI, 436 s. Chapman & Hall, Boca Raton, FL.
- Elliott, J., 1994 *Quantitative Ecology and the Brown Trout*. Oxford University Press, Oxford, New York, Tokyo.
- El Mousadik, A., and R. J. Petit, 1996 High level of genetic differentiation for allelic richness among populations of the argan tree *Argania spinosa* (L) Skeels endemic to Morocco. *Theor. Appl. Genet.* 92: 832–839.
- Evanno, G., S. Regnaut, and J. Goudet, 2005 Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14: 2611–2620.
- Excoffier, L., G. Laval, and S. Schneider, 2007 Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform.* 1: 47–50.
- Fraser, D. J., M. M. Hansen, S. Ostergaard, N. Tessier, M. Legault *et al.*, 2007 Comparative estimation of effective population sizes and temporal gene flow in two contrasting population systems. *Mol. Ecol.* 16: 3866–3889.
- Hansen, M. M., E. E. Nielsen, and K.-L. D. Mensberg, 1997 The problem of sampling families rather than populations: relatedness among individuals in samples of juvenile brown trout *Salmo trutta* L. *Mol. Ecol.* 6: 469–474.
- Hare, M. P., L. Nunney, M. K. Schwartz, D. E. Ruzzante, M. Burford *et al.*, 2011 Understanding and estimating effective population size for practical application in marine species management. *Conserv. Biol.* 25: 438–449.
- Hedrick, P. W., and M. E. Gilpin, 1997 *Genetic Effective Size of a Metapopulation*. Academic Press, New York/London/San Diego.
- Heggnes, J., K. H. Roed, P. E. Jorde, and A. Brabrand, 2009 Dynamic micro-geographic and temporal genetic diversity in vertebrates: the case of lake-spawning populations of brown trout (*Salmo trutta*). *Mol. Ecol.* 18: 1100–1111.

- Hill, W. G., 1972 Effective size of populations with overlapping generations. *Theor. Popul. Biol.* 3: 278–289.
- Hill, W. G., 1979 Note on effective population size with overlapping generations. *Genetics* 92: 317–322.
- Hill, W. G., 1981 Estimation of effective population size from data on linkage disequilibrium. *Genet. Res.* 38: 209–216.
- Jorde, P. E., 2012 Allele frequency covariance among cohorts and its use in estimating effective size of age-structured populations. *Mol. Ecol. Resour.* 12: 476–480.
- Jorde, P. E., and N. Ryman, 1995 Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics* 139: 1077–1090.
- Jorde, P. E., and N. Ryman, 1996 Demographic genetics of brown trout (*Salmo trutta*) and estimation of effective population size from temporal change of allele frequencies. *Genetics* 143: 1369–1381.
- Jorde, P. E., and N. Ryman, 2007 Unbiased estimator for genetic drift and effective population size. *Genetics* 177: 927–935.
- Marshall, G. T. H., A. R. Beaumont, and R. Wyatt, 1992 Genetics of brown trout *Salmo trutta* L. stocks above and below impassable falls in the Conwy River system North Wales. *Aquat. Living Resour.* 5: 9–13.
- Marshall, T. C., J. Slate, L. E. B. Kruuk, and J. M. Pemberton, 1998 Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* 7: 639–655.
- Nei, M., and F. Tajima, 1981 Genetic drift and estimation of effective population size. *Genetics* 98: 625–640.
- Nunney, L., 1999 The effective size of a hierarchically structured population. *Evolution* 53: 1–10.
- Nunney, L., 2000 The limits to knowledge in conservation genetics—the value of effective population size, pp. 179–194 in *Evolutionary Biology*, Vol 32. Kluwer Academic/Plenum, New York.
- Palstra, F. P., and D. E. Ruzzante, 2008 Genetic estimates of contemporary effective population size: What can they tell us about the importance of genetic stochasticity for wild population persistence? *Mol. Ecol.* 17: 3428–3447.
- Palstra, F. P., and D. E. Ruzzante, 2011 Demographic and genetic factors shaping contemporary metapopulation effective size and its empirical estimation in salmonid fish. *Heredity* 107: 444–455.
- Piry, S., A. Alapetite, J. M. Cornuet, D. Paetkau, L. Baudouin *et al.*, 2004 GENECLASS2: a software for genetic assignment and first-generation migrant detection. *J. Hered.* 95: 536–539.
- Pollak, E., 1983 A new method for estimating the effective population size from allele frequency. *Genetics* 104: 531–548.
- Pritchard, J. K., M. Stephens, and P. Donnelly, 2000 Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- Prodohl, P. A., A. F. Walker, R. Hynes, J. B. Taggart, and A. Ferguson, 1997 Genetically monomorphic brown trout (*Salmo trutta* L) populations, as revealed by mitochondrial DNA, multilocus and single-locus minisatellite (VNTR) analyses. *Heredity* 79: 208–213.
- Rannala, B., and J. L. Mountain, 1997 Detecting immigration by using multilocus genotypes. *Proc. Natl. Acad. Sci. USA* 94: 9197–9201.
- Ryman, N., 1983 Patterns of distribution of biochemical genetic variation in salmonids: differences between species. *Aquaculture* 33: 1–21.
- Ryman, N., F. W. Allendorf, and G. Stahl, 1979 Reproductive isolation with little genetic divergence in sympatric populations of brown trout (*Salmo trutta*). *Genetics* 92: 247–262.
- Serbezov, D., L. Bernatchez, E. M. Olsen, and L. A. Vøllestad, 2010 Mating patterns and determinants of individual reproductive success in brown trout (*Salmo trutta*) revealed by parentage analysis of an entire stream living population. *Mol. Ecol.* 19: 3193–3205.
- Serbezov, D., P. E. Jorde, L. Bernatchez, E. M. Olsen, and L. A. Vøllestad, 2012 Life history and demographic determinants of effective/census size ratios as exemplified by brown trout (*Salmo trutta*). *Evol. Appl.* DOI: 10.1111/j.1752-4571.2012.00239.x.
- Tallmon, D. A., G. Luikart, and M. A. Beaumont, 2004 Comparative evaluation of a new effective population size estimator based on approximate Bayesian computation. *Genetics* 167: 977–988.
- Taugbøl, A., 2008 Fine-scale genetic structure of brown trout (*Salmo trutta*). Master's Thesis. University of Oslo, Oslo.
- Van Oosterhout, C., W. F. Hutchinson, D. P. M. Wills, and P. Shipley, 2004 Micro-Checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4: 535–538.
- Wang, J. L., 2009 A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Mol. Ecol.* 18: 2148–2164.
- Wang, J. L., and M. C. Whitlock, 2003 Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* 163: 429–446.
- Waples, R. S., 1989 Generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* 121: 379–391.
- Waples, R. S., 1990 Conservation genetics of Pacific salmon. III. Estimating effective population size. *J. Hered.* 81: 277–289.
- Waples, R. S., and C. Do, 2008 LDNE: a program for estimating effective population size from data on linkage disequilibrium. *Mol. Ecol. Resour.* 8: 753–756.
- Waples, R. S., and P. R. England, 2011 Estimating contemporary effective population size on the basis of linkage disequilibrium in the face of migration. *Genetics* 189: 633–644.
- Waples, R. S., and M. Yokota, 2007 Temporal estimates of effective population size in species with overlapping generations. *Genetics* 175: 219–233.
- Waples, R. S., C. Do, and J. Chopelet, 2011 Calculating Ne and Ne/N in age-structured populations: a hybrid Felsenstein-Hill approach. *Ecology* 92: 1513–1522.
- Wilson, G. A., and B. Rannala, 2003 Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163: 1177–1191.

Communicating editor: M. A. Beaumont