



Scottish Government
Riaghaltas na h-Alba
gov.scot

Using genetic approaches to estimate population sizes of salmon in Scotland

Scottish Marine and Freshwater Science Vol 8 No 4

J Gilbey and P J Bacon



marinescotland

Using genetic approaches to estimate population sizes of salmon in Scotland

Scottish Marine and Freshwater Science Vol 8 No 4

J Gilbey and P J Bacon

Published by Marine Scotland Science

ISSN: 2043-7722

DOI: 10.7489/1893-1

Marine Scotland is the directorate of the Scottish Government responsible for the integrated management of Scotland's seas. Marine Scotland Science (formerly Fisheries Research Services) provides expert scientific and technical advice on marine and fisheries issues. Scottish Marine and Freshwater Science is a series of reports that publishes results of research and monitoring carried out by Marine Scotland Science. It also publishes the results of marine and freshwater scientific work that has been carried out for Marine Scotland under external commission. These reports are not subject to formal external peer-review.

This report presents the results of marine and freshwater scientific work carried out by Marine Scotland Science.

© Crown copyright 2017

You may re-use this information (excluding logos and images) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence, visit: <http://www.nationalarchives.gov.uk/doc/open-governmentlicence/version/3/> or email: psi@nationalarchives.gsi.gov.uk.

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

Using genetic approaches to estimate population sizes of Salmon in Scotland

John Gilbey and Philip J Bacon

Marine Scotland Science Freshwater Fisheries Laboratory

Faskally, Pitlochry, Perthshire, PH16 5LB

Outline

There is considerable interest in the use of genetic tools to assess the conservation status of Atlantic salmon. Here we assess the feasibility of using genetic methods to estimate effective population size and number of breeders, outlining their merits, together with potential drawbacks.

Effective population size (N_e)

Background

In simple terms the effective population size (N_e) can be thought of as the number of reproducing (breeding) individuals in a population, when examined over a single generation. This is a subset of N_c , the adult census population size, defined as the total number of potential (sexually mature) breeders. It is often the case that N_c is unknown as census counts are hard to carry out for all populations in the wild, as such there is a potential to use genetic sampling of fish to estimate N_e which could then act as a surrogate for N_c when accessing the health of populations and making management decisions.

N_e is formally defined as the size of an idealised population that would have the same dispersion of allele frequencies under random genetic drift (or inbreeding) as the observed population (Wright, 1931). It is a key parameter in conservation and management because it affects the degree to which a population can respond to selection and is reflective of both population size and levels of inbreeding. N_e influences the rate of loss of genetic diversity, the rate of fixation of deleterious alleles and the efficiency of natural selection at maintaining beneficial alleles (Berthier et al., 2002). If N_e declines too far, the loss of genetic variation resulting

from genetic drift may put species or populations at risk of extinction by losing the raw material on which selection can operate.

Techniques of estimation

A number of approaches to estimating N_e exist and some of the most used are detailed below. The different estimators can be difficult to compare because different N_e concepts and estimators refer to different time frames and spatial scales and have been applied to many different measures of genetic change. Perhaps the most used variants of N_e are variance N_{eV} , inbreeding N_{eI} and coalescent effective size N_{eC} . Each of these is based on different measures of genetic change, have different assumptions and thus each will estimate a different value for N_e . For overview of metrics see Supplementary Table 1.

A variety of methods have been developed to estimate N_e (see Wang and Whitlock, 2003; Luikart et al., 2010 and references within) using either single sample points in time or multiple temporally-separated samples.

1) Single point samples

Estimation of N_e based on linkage disequilibrium between genetic markers is potentially useful because, unlike most other genetic methods, it requires only a single population sample. The method relies on the fact that in a system where gametes are distributed at random among a small number of zygotes there will be departures from expected genotype frequencies, and departures from expected gametic frequencies, both of which can be used to estimate N_e (England et al., 2006).

2) Temporal samples

A number of approaches can be utilised if temporal sampling points are available. These include examination of short-term allelic frequency changes between sampling periods where N_e estimation is based on comparing observed allele frequencies with those expected changes due to genetic drift between sample times (e.g. Berthier et al., 2002) and coalescent methods where N_e estimation is based on a framework which defines a tree linking the alleles up to their common ancestor and hence describes the relationships among alleles. A coalescence event appears each

time two lineages in the tree join into a common ancestor, and the intervals between such events have a distribution that depends on N_e (Kuhner et al., 1995).

The techniques outlined above give only a brief outline of the possible approaches to N_e estimation. For a more detailed outline see Supplementary Table 2.

The various approaches to estimating N_e have specific assumptions associated with them which typically include that the population is stable, panmictic, including random mating, and there is no selection, migration or mutation (England et al., 2006). Whilst some of these assumptions are likely to be relatively robust for salmon, others will be violated and thus have the potential to cause significant bias to any estimates produced. Of note in this regard is the influence of migrants (both to and from the study site/population) which for salmon populations can substantially bias estimates of N_e if it is not accounted for (Wang & Whitlock et al. 2003). The frequency of matings by precociously mature male parr, which are likely to vary from site to site (Perrier et al. 2014), and the overlapping generations of salmon introduce further complexities (Waples et al., 2014). For detailed overview of assumptions see Supplementary Table 3.

Issues

The variety of theoretical methods available, assumptions associated with them, and statistical approaches available for analysis, together with a situation where models of analysis are developed but often not robustly ground-truthed in real-world situations, means that development of a reliable tool for consistently estimating N_e in a range of Scottish locations is not straight forward. While it is true that genetic data could be collected from Scottish populations and plugged into one or more of the numerous software packages available, and estimates of N_e obtained, there are several pit-falls. For example: the collection of samples (within-years or within generations; across-years or across generations); the actual technique and/or statistical package to use (given the generation composition in relation to the analysis assumptions); the frequency of precocious parr, the presence of migrants, etc, etc. In short, to obtain demonstrably reliable results such approaches to estimating N_e need to be ground-truthed with rarely-available and unusually detailed data for real, wild Scottish populations. This problem is well illustrated by the range

of N_e/N_c ratios which have been estimated for salmonid fish where a review of 98 such estimates across populations within five species had a fourfold to 100-fold difference in ratio estimates (Jones and Avise, 1997). An overview of some major issues that need to be taken into consideration is shown in Fig 1.

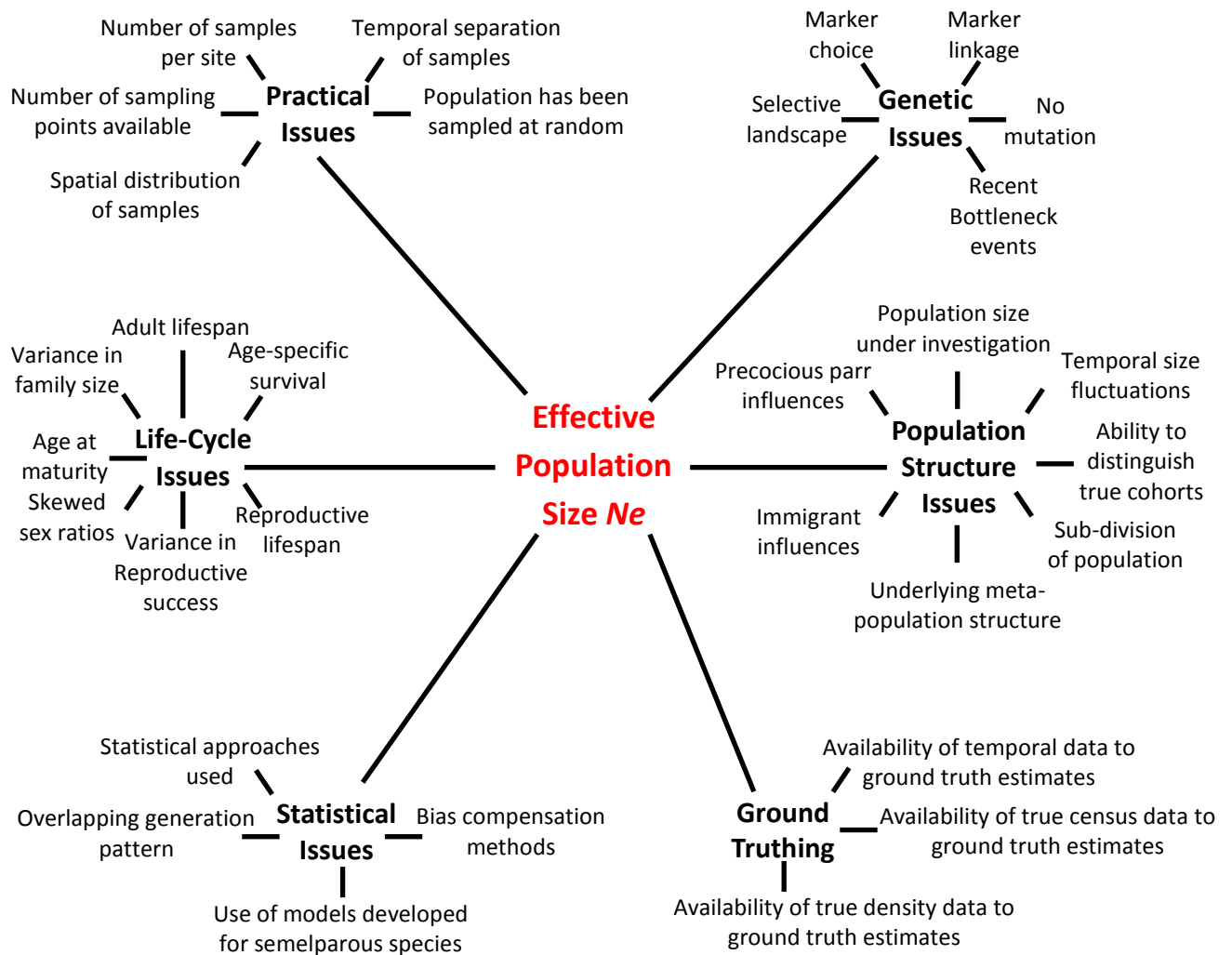


Fig 1. Overview of a selection of issues which need to be considered when calculating N_e .

Ideally sites would be available where genetic estimates of N_e can be compared to more conventional estimates and some of the supporting information required by the different estimators is available. In addition such sites would allow information to be collected on the necessary number of tissue-samples, replicate sampling-sites, and sampling frequencies (across years) required to produce interpretable results. One of the few potential sites where such data exists is the Marine Scotland facility on the Girnock burn, a tributary of the river Dee (described in Creel and Rosenblatt,

2013) (www.gov.scot/Topics/marine/Salmon-Trout-Coarse/Freshwater/Monitoring/Traps), which has recorded various population parameters since the 1960s, and allows a detailed understanding of many of the demographic processes underlying the various N_e metrics. It may also be possible to use long term reconstruction of breeding pedigrees at this site to investigate population size. Initial work is being undertaken to investigate the options for such work at the Girnock. While such work could be crucial to finding an effective method for wider generalisation of a reliable and robust approach for a range of Scottish salmon environments, it would also need to be tested in different situations.

Effective number of breeders (N_b)

The effective number of breeders (N_b) is number of breeders during a single breeding event (Waples et al., 2014). This metric is therefore directly related to N_e because N_b times the generation time approximates N_e (Waples, 1989). To estimate N_b only one temporal sample is required and as such it can be inferred from/for a single cohort (if cohorts can be readily distinguished), and might therefore be an accessible parameter for managers dealing with yearly conservation decisions (Ferchaud et al., 2016).

As with N_e there are a number of statistical techniques which can be utilised to estimate N_b , and again as with N_e , each of these comes with its own set of assumptions and associated potential issues which need to be addressed before robust estimates of N_b can be obtained. The issues under consideration have been outlined above, but in general however, due largely to the fact that N_b is a single cohort time-point estimate, some of the issues associated with the estimation of N_e do not influence N_b calculations and so more robust estimates of N_b can more readily be obtained. This does not however mean that significant problems of interpretation are not still evident. N_b itself may well be a quantitative measure of the number of breeders relating to a sampled breeding event, but how the estimated N_b value relates to the true N_c can still be far from simple to determine, especially if temporal trends need to be examined. Again then, robust ground truthing against known datasets such as that at the Girnock burn, or at sites where quantitative multi-

pass electrofishing has been regularly carried out, are vital to show the true linkage between N_b and N_c at single points and over time.

There are two main approaches to estimation of N_b , Sibship assignment (Wang, 2009) and Linkage Disequilibrium (LD) (Waples and Do, 2008). The Sibship assignment method uses Sibship frequencies estimated from randomly sampled pairs of individuals as being sibs sharing one or two parents. Polygamous breeding of both sexes is allowed for and information on candidate parents of Sibship sizes is not required. The LD method tests for non-random associations between unlinked loci within the dataset (and so again is based around the relatedness of individuals) with a correction due to bias introduced due to the overlapping generations of salmon (Waples et al., 2014).

Relationships between N_e , N_b and N_c

Of critical importance when using genetic methods to examine population size and size trends is how the estimates N_e and N_b relate to the true N_c . The relationships may be true quantitative ones where estimates exactly correlate with census numbers, or qualitative estimates where, even if N_c is not possible to estimate, a stable metric may be obtained which will allow temporal trends to be examined. As before, it is vital therefore, that robust ground truthing is carried out against known census or density data otherwise wildly differing and variable estimations may be wrongly relied upon for important management decisions.

It is thus timely that an important piece of work has been recently undertaken to look at just this issue in Atlantic salmon. Ferchaud et al. (2016) carried out a study focused on “Making sense of the relationships between N_e , N_b and N_c towards defining conservation thresholds in Atlantic salmon (*Salmo salar*)”. They investigated the relationships between N_e , N_b and N_c in 10 Atlantic salmon populations in Québec, Canada, for which they genotyped 100 randomly sampled young-of-the-year individuals for 5 consecutive years. The results showed a positive correlation between N_e , N_b and N_c , suggesting that N_b was an indicative parameter for tracking effective population size and abundance of Atlantic salmon. However, only 27% of

the variation in N_c was explained using estimates of N_b , with a large proportion of the variance in N_b/N_c existing both among populations (37%) and among years (19%). As they go on to say, their results again illustrate the need for thorough calibration of $N_e/N_b/N_c$ before using genetic approaches in monitoring programs, as well as a full understanding of the limits of such approaches.

Genetic mark-recapture (GMR)

With regard to using genetic techniques in estimating population sizes as described above, both N_e and N_b can be used. Genetic markers do however provide a further set of methods to allow investigation of population sizes based on so-called 'Genetic Mark-Recapture' (GMR) estimation, of which again there are a number of possible approaches.

At the most basic level, GMR can be utilised as a simple replacement for external tags utilised for many years in classical mark-recapture programmes. Typically such investigations have been used in situations where population sizes are very small and the physical tagging of individuals is problematic (Lukacs and Burnham, 2005). In such circumstances DNA can be collected from sources such as scats, shed hair or sloughed skin. Each genotype is then considered a 'mark', and a 'recapture' is recorded whenever an identical genotype is found in two separate DNA samples with population size then being estimated using classical mark-recapture algorithms (Mills et al., 2000).

Of perhaps more usefulness to examine population size in salmon are techniques of GMR which utilise genotypes of sampled individuals together with inferences on parentage, relatedness and/or pedigree's to estimate population sizes. For example, recently GMR has been used in novel ways to estimate population size in salmonids (Hamazaki and DeCovich, 2014; Rawding et al., 2014). Such techniques have been used to estimate stock-specific run sizes, escapements, and exploitation rates can also be estimated, potentially providing more information than conventional mark recapture approaches (Hamazaki and DeCovich, 2014).

As with the other techniques listed above there are both practical and technical issues and assumptions that must be considered before such approaches could be considered for stock estimation in novel Scottish situations. Robust ground-truthing would be required in order to produce a workable tool.

Summary

Effective fishery management benefits greatly from on an understanding of both population ecology and genetics. In turn, this understanding requires robust information about population size and dynamics, distribution patterns and limits, reproductive strategy, and ability to adapt to abiotic and biotic changes. Perhaps the primary factor of import to this process is the size of the population, one of the fundamental parameters in both fishery management and evolutionary biology (Hare et al., 2011). In many cases however, this simple metric is very hard to collect. Rivers contain many sub-structured genetic populations of Atlantic salmon, each interacting with one another to a greater or lesser extent. Obtaining reliable census counts of these myriad sub-populations is likely to always be impossible. In such a situation genetic tools provide great promise in their ability to be able to provide robust estimates of parameters related to sub-population sizes with which conservation status can be assessed at a local stock/population level and informed management decisions made. As such Marine Scotland Science is continuing to investigate the development of reliable methods and to avoid/negate the potential pitfalls which may occur if robust ground truthing of techniques is not carried out.

References

- Anderson, E. C. 2005. An Efficient Monte Carlo Method for Estimating N_e From Temporally Spaced Samples Using a Coalescent-Based Likelihood. *Genetics*, 170: 955-967.
- Anderson, E. C., Williamson, E. G., and Thompson, E. A. 2000. Monte Carlo Evaluation of the Likelihood for N_e From Temporally Spaced Samples. *Genetics*, 156: 2109-2118.
- Balloux, F. 2004. Heterozygote excess in small populations and the heterozygote-excess effective population size. *Evolution*, 58: 1891-1900.

- Beaumont, M. A. 2003. Estimation of Population Growth or Decline in Genetically Monitored Populations. *Genetics*, 164: 1139-1160.
- Berthier, P., Beaumont, M. A., Cornuet, J.-M., and Luikart, G. 2002. Likelihood-Based Estimation of the Effective Population Size Using Temporal Changes in Allele Frequencies: A Genealogical Approach. *Genetics*, 160: 741-751.
- Creel, S., and Rosenblatt, E. 2013. Using pedigree reconstruction to estimate population size: genotypes are more than individually unique marks. *Ecology and Evolution*, 3: 1294-1304.
- Crow, J. F., and Denniston, C. 1988. Inbreeding and Variance Effective Population Numbers. *Evolution*, 42: 482-495.
- England, P. R., Cornuet, J.-M., Berthier, P., Tallmon, D. A., and Luikart, G. 2006. Estimating effective population size from linkage disequilibrium: severe bias in small samples. *Conservation Genetics*, 7: 303-308.
- Ewens, W. J. 1982. On the concept of the effective population size. *Theoretical Population Biology*, 21: 373-378.
- Ferchaud, A. L., Perrier, C., April, J., Hernandez, C., Dionne, M., and Bernatchez, L. 2016. Making sense of the relationships between N_e , N_b and N_c towards defining conservation thresholds in Atlantic salmon (*Salmo salar*). *Heredity*.
- Hamazaki, T., and DeCovich, N. 2014. Application of the Genetic Mark–Recapture Technique for Run Size Estimation of Yukon River Chinook Salmon. *North American Journal of Fisheries Management*, 34: 276-286.
- Hare, M. P., Nunney, L., Schwartz, M. K., Ruzzante, D. E., Burford, M., Waples, R. S., Ruegg, K., et al. 2011. Understanding and estimating effective population size for practical application in marine species management. *Conserv Biol*, 25: 438-449.
- Harris, R. B., and Allendorf, F. W. 1989. Genetically Effective Population Size of Large Mammals: An Assessment of Estimators. *Conservation Biology*, 3: 181-191.
- Hauser, L., Adcock, G. J., Smith, P. J., Bernal Ramírez, J. H., and Carvalho, G. R. 2002. Loss of microsatellite diversity and low effective population size in an overexploited population of New Zealand snapper (*Pagrus auratus*). *Proceedings of the National Academy of Sciences*, 99: 11742-11747.
- Hedgecock, D., Launey, S., Pudovkin, A. I., Naciri, Y., Lapègue, S., and Bonhomme, F. 2006. Small effective number of parents (N_b) inferred for a naturally spawned cohort of juvenile European flat oysters *Ostrea edulis*. *Marine Biology*, 150: 1173.
- Jones, A. G., and Avise, J. C. 1997. Microsatellite analysis of maternity and the mating system in the Gulf pipefish *Syngnathus scovelli*, a species with male pregnancy and sex-role reversal. *Molecular Ecology*, 6: 203-213.
- Jorde, P. E., and Ryman, N. 1995. Temporal allele frequency change and estimation of effective size in populations with overlapping generations. *Genetics*, 139: 1077-1090.
- Jorde, P. E., and Ryman, N. 2007. Unbiased estimator for genetic drift and effective population size. *Genetics*, 177: 927-935.
- Kuhner, M. K., Yamato, J., and Felsenstein, J. 1995. Estimating effective population size and mutation rate from sequence data using Metropolis-Hastings sampling. *Genetics*, 140: 1421-1430.
- Luikart, G., Ryman, N., Tallmon, D. A., Schwartz, M. K., and Allendorf, F. W. 2010. Estimation of census and effective population sizes: the increasing usefulness of DNA-based approaches. *Conservation Genetics*, 11: 355-373.

- Lukacs, P. M., and Burnham, K. P. 2005. Review of capture–recapture methods applicable to noninvasive genetic sampling. *Molecular Ecology*, 14: 3909-3919.
- Miller, C. R., and Waits, L. P. 2003. The history of effective population size and genetic diversity in the Yellowstone grizzly (*Ursus arctos*): implications for conservation. *Proceedings of the National Academy of Science, USA.*, 100: 4334-4339.
- Mills, L. S., Citta, J. J., Lair, K. P., Schwartz, M. K., and Tallmon, D. A. 2000. Estimating animal abundance using non-invasive DNA sampling: promise and pitfalls. *Ecological Applications*, 10: 283-294.
- Nomura, T. 2008. Estimation of effective number of breeders from molecular coancestry of single cohort sample. *Evolutionary Applications*, 1: 462-474.
- Ovenden, J. R., Peel, D., Street, R., Courtney, A. J., Hoyle, S. D., Peel, S. L., and Podlich, H. 2007. The genetic effective and adult census size of an Australian population of tiger prawns (*Penaeus esculentus*). *Molecular Ecology*, 16: 127-138.
- Peel, D., Ovenden, J. R., and Peel, S. L. 2004. NeEstimator: software for estimating effective population size, Version 1.3. Queensland Government, Department of Primary Industries and Fisheries.
- Rawding, D. J., Sharpe, C. S., and Blankenship, S. M. 2014. Genetic-Based Estimates of Adult Chinook Salmon Spawner Abundance from Carcass Surveys and Juvenile Out-Migrant Traps. *Transactions of the American Fisheries Society*, 143: 55-67.
- Tallmon, D. A., Koyuk, A., Luikart, G., and Beaumont, M. A. 2008. ONeSAMP: a program to estimate effective population size using approximate Bayesian computation. *Molecular Ecology Resources*, 8: 299-301.
- Vitalis, R., and Couvet, D. 2001. Estimation of effective population size and migration rate from one- and two-locus identity measures. *Genetics*, 157: 911-925.
- Wakeley, J., and Sargsyan, O. 2009. Extensions of the Coalescent Effective Population Size. *Genetics*, 181: 341-345.
- Wang, J. 2001. A pseudo-likelihood method for estimating effective population size from temporally spaced samples. *Genet Res*, 78: 243-257.
- Wang, J. 2009. A new method for estimating effective population sizes from a single sample of multilocus genotypes. *Molecular Ecology*, 18: 2148-2164.
- Wang, J., and Whitlock, M. C. 2003. Estimating effective population size and migration rates from genetic samples over space and time. *Genetics*, 163: 429-446.
- Waples, R. S. 1989. A 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. *Journal of Heredity*, 81: 277-289.
- Waples, R. S. 2002. Definition and estimation of effective population size in the conservation of endangered species. *In Population viability analysis*, pp. 147-168. Ed. by S. R. Beissinger, and D. R. McCullough. The University of Chicago Press, Chicago.
- Waples, R. S. 2005. Genetic estimates of contemporary effective population size: to what time periods do the estimates apply? *Molecular Ecology*, 14: 3335-3352.
- Waples, R. S., Antao, T., and Luikart, G. 2014. Effects of Overlapping Generations on Linkage Disequilibrium Estimates of Effective Population Size. *Genetics*.

- Waples, R. S., and Do, C. 2010. Linkage disequilibrium estimates of contemporary $N(e)$ using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, 3: 244-262.
- Waples, R. S., and Do, C. H. I. 2008. Idne: a program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, 8: 753-756.
- Wright, S. 1931. Evolution in Mendelian Populations. *Genetics*, 16: 97-159.
- Wright, S. 1969. *Evolution and the Genetics of Populations, Volume 2. Theory of Gene Frequencies*, University of Chicago Press, Chicago.
- Zhdanova, O. L., and Pudovkin, A. I. 2008. Nb_HetEx: A Program to Estimate the Effective Number of Breeders. *Journal of Heredity*, 99: 694-695.

Supplementary Table 1. Kinds of N_e estimators and concepts.

Effective population size is whatever must be substituted in the formula $(1/2N)$ to describe the actual loss in heterozygosity (Wright, 1969)

There are a number of ways to estimate contemporary N_e which each fall into one of two categories; direct demographic methods or indirect genetic methods. Care must be taken when using any of the possible methods due to failure to include all influencing factors which can change the relationship between N_e compared to the true N_c .

It can be difficult to compare different estimates of N_e due to the many different measures of genetic change they have been applied to. N_{eV} and N_{eI} are the most widely used estimators in conservation and management and the best evaluated. However, even these are difficult to compare as the inbreeding N_e (N_{eI}) is concerned with the loss of heterozygosity while the variance N_e (N_{eV}) is concerned with change in allele frequencies through time. A number of other forms of N_e also exist (Ewens, 1982; Crow and Denniston, 1988; Wakeley and Sargsyan, 2009).

In stable populations the inbreeding N_e (N_{eI}) and variance N_e (N_{eV}) are similar and distinguishing between them is unimportant for conservation purposes (Waples, 2002). However, populations of conservation interest are often changing rapidly in size, or have been through extreme events such as bottlenecks and in such situations the two measures may differ markedly. In a declining population N_{eI} can be considerable higher than N_{eV} for a considerable period of time whereas in a recovering population N_{eV} will recover faster than N_{eI} to previous levels. Thus, it can be difficult to understand the practical consequences of the N_{eV} vs N_{eI} distinction because they can refer to different time periods depending on if parents and/or offspring are sampled and which N_e estimator is used (Waples, 2005).

The coalescent effective size (N_{eC}) concept considers, in theory, all aspects of genetic change, whereas other forms of N_e (N_{eV} and N_{eI}) include only a single measure of the rate of genetic drift (variance in allele frequencies) or inbreeding (heterozygosity). Thus the coalescent N_e might sometimes be preferable because the coalescent holds for a surprisingly wide range of population models including the Wright-Fisher models. Coalescent-based N_e estimators also perform well in small populations (Anderson, 2005) even though the assumption of only one coalescent event per generation is likely violated in such populations.

Together with difficulties in estimating contemporary N_e and comparing between different methods, it should also be mentioned that significant difficulty is also evident when measuring long-term N_e . Depending on the values defined for the N_e models being used, very different estimates may be forthcoming. For example Ovenden et al. (2007) estimated that long-term N_e was 10 fold higher than the contemporary N_e in tiger prawns (*Penaeus esculentus*) from Morton Bay, Australia. However as the authors point out, this could simply be a result arising from assuming a mutation rate of 10^{-3} rather than 10^{-4} ; the lower mutation rate would require a higher N_e to yield the same heterozygosity.

Supp Table 1 after Box 1 of Luikart et al. (2010). © Springer Science+Business Media B.V. 2010, reproduced with permission of Springer

Supplementary Table 2. Genetic estimators of contemporary N_e

N_e estimator	Strengths	Key assumptions	Software and references
<u>One sample</u>			
Linkage Disequilibrium (LD)	Uses any ~10–20 unlinked loci, and 30–50 individuals	LD signal arises only from genetic drift	<i>LD-Ne</i> ; Waples and Do (2008; 2010)
Approximate Bayesian method using LD (plus 7 other summary statistics)	Uses more information than the LD method; Allows prior on N_e	LD signal arises only from genetic drift effects	<i>ONE-SAMP</i> ; Tallmon et al. (2008)
Heterozygote excess	Estimates N_b from single sample if N_b is very small	Signal only from different allele frequencies in male & female breeders	<i>Nb_HetEx</i> ; Zhdanova and Pudovkin (2008); Balloux (2004)
Identity dis-equilibrium at 1 & 2 loci	Estimates N_e and migration rate jointly	LD signal is from genetic drift and migration	Vitalis and Couvet (2001)
Molecular coancestry (i.e. allele sharing among sampled individuals)	Estimates N_b from single sample if N_b is very small	Non-sib pairs Needed as reference for co-ancestry among individuals	Nomura (2008)
Sib identification	Applies to non-random mating populations, codominant & dominant loci	Sibs & relatedness are reliably identified. No/low immigration	<i>Colony2</i> ; Wang (2009)
Rarefaction of alleles	Estimates of N_b ; precision similar to the temporal method	Progeny are produced from few adults in a large H–W equilibrium population	Hedgewood et al. (2006)
<u>Two samples</u>			
Heterozygosity decline	Computation is simple; much theory behind heterozygosity	Decrease in heterozygosity is caused only by small N_e	Harris and Allendorf (1989); Hauser et al. (2002); Miller and Waits (2003)
Temporal F-statistic moments method	Computationally rapid	Allele frequency change is only from drift; No selection or migration	<i>Ne-estimator</i> ; Peel et al. (2004); <i>TempoFs</i> ; Jorde and Ryman (2007) uses unbiased estimator
Pseudo-ML (maximum likelihood) temporal method	Computationally rapid; allows for migration	Allele frequency change arises only from drift (& migration if also estimating m)	<i>MLNE</i> ; Wang (2001), Wang and Whitlock (2003)
ML and MCMC temporal method	Useful on multi-allelic loci	Allele frequency changes only from drift	<i>MCLEEPS</i> ; Anderson (2000)
Coalescent Bayesian temporal method	Allows prior on N_e which can improve precision	Same as just above; One coalescent event per generation	<i>TM3</i> ; Berthier et al. (2002); <i>CoNe</i> ; Anderson (2005)
<u>Three samples</u>			
Coalescent Bayesian	Allows prior on N_e	Same as just above	<i>TMVP</i> ; Beaumont (2003)

Supp Table 2 after Table 3 of Luikart et al. (2010). © Springer Science+Business Media B.V. 2010, reproduced with permission of Springer

Supplementary Table 3. Assumptions common to many N_e estimators, and approaches to avoid violating assumptions.

Assumption	Likelihood & consequences of violating assumptions, and ways to avoid violations
Random sampling carried out	Likely to be violated unless precautions are taken. Should be tested for over-representation of family groups and for spatial structuring (e.g. Hardy–Weinberg tests or clustering).
Random loci choice	Often violated. Highly polymorphic loci often preferred non-random, choice. Tests for linkage should be performed if using 100's or 1,000's of loci with strongly linked loci not being used unless linkage is accounted for. Tests for non-independence should be conducted and non-independent loci excluded (except for LD- N_e which uses inter-locus associations to estimate N_e).
Population is not sub-divided	Likely occasionally or often violated. Can significantly bias populations especially if varying proportions of each sub-population in temporally separate samples. Tests should be carried out for clusters, substructure & Wahlund effects before estimating N_e (e.g. H–W tests or clustering).
No immigration	Likely to be occasionally or often violated. Tests could be performed (e.g. assignment tests) with immigrants removed before estimating N_e . LD- N_e and temporal N_e appear insensitive to limited immigration ($m < 0.10$) in fragmenting populations and the assumption is relaxed in temporal method of Wang and Whitlock (2003).
No mutation	Likely not violated for most loci in most contemporary N_e estimates likely often violated for long-term N_e estimates for which a mutation model and rate must be estimated (assumed).
Selection is negligible	Likely not violated as most loci used for these approaches are effectively neutral. Tests for neutrality and outlier loci should be conducted before estimating N_e or N_e estimated jointly with selection.
No overlapping generations and no age structure	Likely occasionally or often violated in Atlantic salmon sampling. LD- N_e likely biased (e.g. by gametic disequilibrium generated by overlapping generations). The generational composition of juvenile and adult samples could be identified by scale reading and cohorts identified. Assumption relaxed in the modified temporal methods of Jorde and Ryman (1995) and Waples (1990).
Population size is stable	Likely to be often violated. Possibly detectable with bottleneck tests and by comparing N_{ev} and N_{el} estimates. Relaxed in recently developed estimators for the temporal method and coalescent methods and bias-effects poorly understood.

Supp Table 3 after Table 4 of Luikart et al. (2010). © Springer Science+Business Media B.V. 2010, reproduced with permission of Springer