Identification of management units using population genetic data

Per J. Palsbøll¹,², Martine Bérubé¹ and Fred W. Allendorf³,⁴

¹Department of Environmental Science, Policy and Management, University of California Berkeley, Berkeley, California, 94720, USA
²Department of Genetics, Microbiology, and Toxicology, Stockholm University, S-106 91 Stockholm, Sweden
³School of Biological Sciences, Victoria University of Wellington, Wellington, New Zealand
⁴Division of Biological Sciences, University of Montana, Missoula, MT 59812, USA

The identification of management units (MUs) is central to the management of natural populations and is crucial for monitoring the effects of human activity upon species abundance. Here, we propose that the identification of MUs from population genetic data should be based upon the amount of genetic divergence at which populations become demographically independent instead of the current criterion that focuses on rejecting panmixia. MU status should only be assigned when the observed estimate of genetic divergence is significantly greater than a predefined threshold value. We emphasize the need for a demographic interpretation of estimates of genetic divergence given that it is often the dispersal rate of individuals that is the parameter of immediate interest to conservationists rather than the historical amount of gene flow.

Introduction

Management units (MUs; see Glossary) are usually defined as demographically independent populations whose population dynamics (e.g. population growth rate) depend largely on local birth and death rates rather than on immigration. The identification of MUs is central to the short-term management and conservation of natural populations and is typically used to delineate entities for monitoring and regulating the effects of human activity upon the abundance of populations and species. Given that MUs represent demographically isolated units, their delineation requires an estimate of dispersal rates among populations. Over the past decade or so, the use of population analyses of genetic markers has increased substantially as an indirect means of inferring whether subpopulations constitute part of the same MU. The criterion most commonly used to delineate MUs was proposed more than a decade ago by Moritz, who defined MUs as ‘…populations with significant divergence of allele frequencies at nuclear or mitochondrial loci, regardless of the phylogenetic distinctiveness of the alleles…’ [2]. The wording ‘significant divergence’ has since been inferred to mean the statistical rejection of panmixia, and continues to be the yardstick used when designating MU status from population genetic data (Box 1). This interpretation is evident in several recent reviews dealing specifically with the application of population genetic data to wildlife and fisheries management (e.g. Refs [3,4]) as well as in the assessment of the performance of the statistics used to define MUs (e.g. Ref. [5]). Here, we advocate basing the delineation of MUs upon the amount of population genetic divergence instead of simply the rejection of panmixia.

Delineation of MUs should be based upon the observed estimate of population genetic divergence

Focusing on rejecting panmixia rather than on the amount of population genetic divergence could misguide the
Box 1. Management units in Scandinavian brown bear Ursus arctos

A series of papers describing the genetic differentiation of Scandinavian brown bears Ursus arctos provides an excellent example of using genetic information to delineate MUs and the associated issues when the delimiting criterion is not defined a priori. Four subpopulation clusters (NN, NS, M and S, from north to south) were identified by the geographical distribution of bears killed by hunters. mtDNA analysis found two highly divergent haplotypes with a discrete geographical distribution and a contact zone separating the S and three more northern subpopulations [36]. Thus, it was proposed that the southern and three more northern clusters of bears should be treated as two evolutionarily significant units (ESUs).

Subsequent analysis of 19 microsatellite loci in the same individuals reached a different conclusion [37]. The null hypothesis of panmixia was rejected for all pairwise comparisons with a high degree of certainty (P < 0.001). In addition, these nuclear data gave no indication of a break between the S and three northern subpopulations. Rather, FST values between adjacent subpopulations were estimated at ~0.08 for all pairwise comparisons, except between the nearby NN and NS subpopulation pairs, where FST was estimated at 0.015. These authors recommended that these bear populations should be considered a single ESU, but divided into four separate MUs because of the rejection of panmixia in all pairwise comparisons. Subsequently, Manel et al. [38] reanalyzed the data of Waits et al. [36] without using the a priori subpopulation designations based upon geography. They found evidence for three subpopulations that matched the three primary geographical clusters of individuals. The NN and NS subpopulations were combined into a single subpopulation and three MUs were defined (S, M and N). These results demonstrate the importance of determining genetic structure from the data, without presupposing a structure.

Tallmon et al. [39] later determined genotypes at the same 19 microsatellite loci as used above in samples taken ~15 years after the samples used by Waits et al. [36]. These authors were interested primarily in the conservation status of the S subpopulation. They estimated N0 of this subpopulation at ~45, and the number of immigrants into this subpopulation at ~0.5, yielding a dispersal rate of ~1%. They concluded that this subpopulation was demographically isolated, but that the amount of gene flow into it was sufficient to avoid inbreeding depression.

The Scandinavian brown bear studies serve as an excellent (and common) case of how the conclusions (in terms of ESU and MU status) change as different kinds of analysis population genetic assessment are conducted, even when the same data are used. However, in none of the studies was the uncertainty of the degree of genetic divergence estimated. Perhaps more importantly, the delimiting criteria were not defined a priori, and only in the study by Tallmon et al. [39] was the genetic estimate of migration related to the level of demographic (as well as genetic) independence of the target population.

How genetically divergent should different MUs be?

At what amount of population genetic divergence should populations be assigned to different MUs? A ‘one size fits all’ answer is not possible given that it depends upon the specific conservation context, as well as on the biological characteristics and population history of the target species [11,13–15]. More importantly, as Waples and Gaggiotti [7] point out, there is currently no general framework for determining at which dispersal rate populations becomes demographically correlated. Work by Hastings [16] suggests that populations become demographically correlated at dispersal rates above 10%. In such cases, populations should be assigned to different MUs if the rate of dispersal among populations is <10%. If a population genetic approach is used to delineate MUs, then it will be necessary to determine what amount of population genetic divergence corresponds to the relevant dispersal rate (e.g. 10%).
The number of migrants is usually estimated as $mN_e$, which is the product of the effective population size ($N_e$) and migration rate ($m$). The probability that gene allele frequencies distributions in the two samples are collected from the same panmictic population was estimated using a permutation-based $\chi^2$ test [40].

Even though the dispersal rate (here set at 20%) was well above the 10% proposed by Hastings [16] as the level of dispersal at which populations becomes demographically correlated (and thus should be assigned to the same MU), a large proportion of the population genetic assessments resulted in panmixia rejection. For large sample sizes (i.e. 100 or 200 individuals per population), between 10 to 45% of the assessments resulted a panmixia rejection and, hence, an erroneous assignment of the two populations to different MUs.

Using a Bayesian approach

Figure 1b depicts the posterior probability distribution of the parameter $mN_e$ estimated [11] from simulated data from two putative MUs; at five (black line) and 15 (gray line) loci. The vertical dashed line indicates the location of the predefined threshold value of $mN_e$. Because a Bayesian likelihood approach is used, the relative probabilities of one or two MUs (given a threshold value for $mN_e$) can be inferred. In this case, ~89% and ~96% of the posterior probability distribution estimated from five and 15 loci, respectively, is below the threshold value. Hence, although the assessment based upon five loci fail to resolve the MU assignment, the Bayesian likelihood approach used enables a probabilistic statement of each of the two hypotheses; one or two MUs. Here, the probability of $mN_e$ being larger than the threshold value was estimated at 0.11 (one MU); conversely, the probability of the alternate model (two different MUs) is 0.89. This approach is useful when resources are insufficient to collect and analyze additional data to resolve fully the MU status, or immediate management action is called for.

Assuming selective neutrality and equilibrium conditions (referred to as migration–drift equilibrium), the amount of genetic divergence among populations is a function of the number of migrants per generation (also termed ‘gene flow’) rather than the dispersal rate per se. The number of migrants is usually estimated as $mN_e$, which is the product of the effective population size ($N_e$) and the probability that an individual is a migrant ($m$). Consequently, a migration rate of, for instance, 10% will result in a much higher degree of genetic divergence among populations with $N_e$ of 1000 compared with population with $N_e$ of 100. The rate of gene flow is usually scaled to either the $N_e$ (as above) or the mutation rate [11]. Given that the size of these two parameters is usually unknown, it is difficult to translate gene flow into dispersal rates, which is the parameter of interest for determining whether populations are demographically correlated (i.e. part of the same or different MUs). In addition, most of the current approaches used to estimate the degree of genetic divergence among populations (such as Wright’s $F_{ST}$ [17]) use the relative degree of genetic diversity within and among populations to infer $mN_e$. However, as $mN_e$ becomes large (>5–10), statistical uncertainty increases, resulting in either poor statistical precision [9,18,19] or erroneous estimates [20]. This inverse correlation between statistical precision and migration rates implies that it will require comparatively few data to reject high migration rates, as would be the case when the sampled populations constitute multiple MUs.

Finally, dispersal rate and gene flow are two different entities that might differ substantially. An extreme example is the study of North Atlantic pilot whales *Globicephala melanos* by Amos and co-workers [21]. The study revealed that male pilot whales stay with their maternal pod even after they become sexually mature, but mature males do not breed with the females in their resident pod. In this case, the degree of dispersal from the level of genetic divergence among pilot whale pods would be misleading, and dispersal could be zero even at very high levels of gene flow. Dispersal rates might be high among populations, but the rate of gene flow could be much lower if immigrants are incapable of contributing successfully to the local gene pool.

For all these reasons, to define an appropriate threshold level of population genetic divergence at which populations should constitute separate MUs, it will be necessary to establish the relationship between the demographic characteristics and population genetic dynamics of the target species. Given the lack of a general analytical framework that combines the demographic and genetic characteristic of populations and their individuals [7], such correlations are currently best determined using individual-based
Box 3. Defining MUs of sockeye salmon using the revised criterion

Salmon on the west coast of North America are often harvested in mixed stock fisheries that contain fish from many local reproductive populations that are returning to spawn in their natal streams. Fishery managers have long recognized the importance of managing local populations individually so that an adequate number of individuals from each local population is allowed to reproduce to ensure the persistence of the local populations (i.e. MUs) that make up fishery stock [41]. Up to 40 million sockeye salmon *Oncorhynchus nerka* are captured each year in Bristol Bay, Alaska [42]. These salmon are returning to spawn at several hundred natal sites in streams, rivers and lakes that drain into Bristol Bay. Lake Clark is one of two large lake systems in the Kvichak River system, which has historically been the largest contributor of sockeye salmon to the Bristol Bay fishery.

Ramstad *et al.* [43] analyzed ~100 sockeye salmon from 11 spawning sites throughout the Lake Clark drainage at 11 microsatellite loci to determine whether these sites are demographically isolated. Approximately 1–5000 fish spawn in each of these sites annually [44]. Waples [45] showed that, for Pacific salmon, *N*$_0$ is approximately equivalent to the effective number of breeders per year *x* the average age of reproduction (approximately four years for Lake Clark sockeye salmon). In addition, the *N*$_0$ census population size ratio in Pacific salmon is ~0.2–0.46 [46]. Therefore, the *N*$_0$ for each of the Lake Clark spawning sites is ~1000 or slightly greater. Using the criterion of at least 10% exchange [16], these sites would be demographically isolated if they exchanged <100 or so adults. This corresponds to an *F*$_ST$ of 0.0025 under a Wright–Fisher island population model. Therefore, we would conclude that these spawning sites constitute separate MUs if their genetic divergence exceeds *F*$_ST$ = 0.0025.

The overall value of *F*$_ST$ among these sites is ~0.018 (95% CI 0.010–0.029), greater than our threshold of 0.0025. This corresponds to ~15 migrants per generation with the Wright–Fisher island population model. However, much of this divergence results from one sample site that tended to show greater pairwise divergence than the other ten sites. If we exclude this site, the overall value of *F*$_ST$ among Lake Clark sites is ~0.007 (95% CI 0.004–0.010). This is still at least twice as great as our threshold of 0.0025. Therefore, we conclude that these 11 spawning sites are demographically isolated and should be considered separate MUs.

population computer models that enable the joint generation of demographic and genetic data. Currently, there are only a few computer simulation programs [22–24] that combine demographic and genetic modeling, by which the expected level of population genetic divergence might be estimated under a specific demographic model and dispersal rate.

This crucial link between biologically realistic demographic models and population genetic estimation requires considerable development. Current population genetic inferences rely upon highly idealized and simplistic population models that do not apply to most natural populations [25]. It is important to know when the conclusions from these models are robust with regard to deviations from the underlying assumptions. The sensitivity of biologically feasible deviations from the underlying population genetic model should be assessed with computer simulations before making firm recommendations.

What population genetic divergence measure should be used to delineate MUs?

Inferring the rate of gene flow among putative MUs from the level of genetic variation within and among population samples relies heavily upon several highly simplifying model assumptions, many of which are unlikely to be met in natural populations [25]. The most basic assumption is that of migration–drift equilibrium, which assumes constant population sizes and migration rates over recent evolutionary time, as well as non-overlapping generations and the absence of natural selection. Although some recent developments in data analyses of population genetic data enable estimation of *mN*$_e$ under non-equilibrium conditions, these approaches are either limited to pairwise comparisons [26] or ignore the effect of population divergence time and enable only a simple exponential change in population size [27]. Estimates of gene flow can be severely biased owing to gene flow from other populations that are not included in the estimation [28].

Many standard population genetic methods yield estimates of population divergence that are the average over recent evolutionary time, rather than the current rate of gene flow. In defining MUs, the current rate of dispersal is the measure of interest [29]. Several recent experimental and analytical advances in molecular ecology have been directed towards estimating dispersal rates from genetic data that apply to the last few generations. Such ‘real time’ estimates of dispersal could be obtained by genetic tagging and recapture of individuals [30], or by the distribution of close relatives among the collected samples identified from genetic analyses of kinship [31,32]. Population allele frequencies could also be used to assign individuals to specific populations based on assignment tests [33]. Some assignment-based methods divide the collected samples into clusters of individuals, thereby circumventing the need for a subjective *a priori* partition of collected samples. These methods are particularly useful when the overall knowledge regarding the population structure is poor. However, after samples have been partitioned using such an assignment approach, the degree of genetic divergence (and its uncertainty) among the resulting clusters (each assumed to be a separate population) needs to be estimated before a MU status can be assigned to each cluster.

Individual-based assignment methods can identify descendants of immigrants, which will appear to have ‘hybridized genomes’, with one-half of their genome originating from their resident population and the other from the native population of the immigrant parent. Wilson and Rannala [12] recently developed a Bayesian likelihood approach that uses this aspect to estimate recent dispersal rates (*m*) from population samples of multi-locus genotypes.

Finally, there are also populations that are structured in a manner that makes any MU delineation arbitrary. One example is the biologically realistic and probably common ‘Stepping Stone’ population model. In this case, an assessment of genetic connectivity might suggest that the two geographically most distant populations should each be assigned to separate MUs, but that any geographically intervening populations might be assigned to either of the two MUs (Box 1).

Conclusions

The revised procedure proposed here requires a better understanding of the interaction between population
genetic and demographic parameters and their analysis. Until the appropriate analytical framework is developed, the most efficient manner to assess the effect of different demographic parameter values upon the genetic makeup of putative MUs is to use computer simulations.

The procedure that we propose here to delineate MUs is also applicable to non-genetic measures of population connectivity from where a point estimate and its associated uncertainty can be obtained, such as traditional mark–recapture data [34] or even the degree of similarity of morphological traits [35]. The only requirement is that the statistic used in the assessment can be related to the relevant demographic processes (usually dispersal).

Our proposal to define MUs upon the amount of population genetic divergence (as opposed to the ability to reject panmixia) is general, but yet flexible so that the threshold values of population genetic divergence can be tailored specifically to each particular biological and conservation context. Perhaps equally important is that our proposal requires each assessment to define explicitly the demographic assumptions and the manner in which they are translated into a measure of population genetic divergence. The need to define the delimiting criteria a priori makes the process of assigning MU status to the target populations both transparent and explicit, and readily amendable to changes as new insights are obtained.

Acknowledgements
We are indebted to our institutions for supporting us in this endeavor. We are also grateful for the insightful comments on various drafts by our colleagues Zachariah Peery, Clinton Epps, Jeffrey Lozier, as well as Robin Waples and Craig Moritz. Phillip Hedrick and Oscar Gaggiotti (as well as a third anonymous referee) provided useful suggestions for improvements.

References
17 Wright, S. (1931) Evolution in mendelian populations. Genetics 16, 97–159
41 Rich, W.H. (1939) Local Populations and Migration in Relation to the Conservation of Pacific Salmon in the Western States and Alaska, Contributions No. 1, Fish Commission of Oregon

Forthcoming Conferences

Are you organizing a conference, workshop or meeting that would be of interest to TREE readers? If so, please e-mail the details to us at TREE@elsevier.com and we will feature it in our Forthcoming Conference filler.

19–24 February 2007
International Conference on the Environment: Survival and Sustainability, Nicosia, Cyprus
http://www.neuconference.org/

1–4 March, 2007
Environmental History Meeting, Baton Rouge, LO, USA
http://www.lsu.edu/lsupress/catalog/Fall2004/books/Colten_Uunnatural.html

15–20 April 2007
The 3rd Global Botanic Gardens Congress, Wuhan, PR China
http://www.3gbgc.com

13–19 May 2007
VIIIth Neotropical Ornithological Congress, Estado Monagas, Venezuela
(http://www.esa.org/federation/meetings/FirstCallContributions_VIIIINOC.pdf)

22–27 May 2007
EcoSummit 2007: Ecological Complexity and Sustainability: Challenges and Opportunities for 21st-Century’s Ecology, Beijing, PR China
http://www.ecosummit2007.elsevier.com

5–8 June 2007
ESEE 2007. Integrating Natural and Social Sciences for Sustainability, Leipzig, Germany
http://www.esee2007.ufz.de

16–21 June 2007
Evolution Annual Meeting, held jointly by the Society for the Study of Evolution, the American Society of Naturalists and the Society of Systematic Biologists, Christchurch, New Zealand
http://www.evolutionsociety.org/meetings.htm

1–5 July 2007
Society for Conservation Biology Annual Meeting, Port Elizabeth, South Africa
http://www.nmmu.ac.za/scb/

16–18 July 2007
ENTO ’07: RES Annual National Meeting and RES Symposium on Aquatic Insects, Edinburgh, UK
http://www.royensoc.co.uk/

5–10 August 2007
92nd ESA Annual Meeting, held jointly with SER, San Jose, CA, USA
http://www.esa.org/meetings/FutureAnnualMeetings.php

12–18 August 2007
30th Congress of the International Association of Theoretical and Applied Limnology, Montréal, Canada
(http://www.sil2007.org)

20–25 August 2007
11th Congress of The European Society for Evolutionary Biology, Uppsala, Sweden
(http://www.eseb.org/)