Updated but still under development: Guidelines for the analysis of population genetic data used in an IWC management context

ROBIN S. WAPLES[1], A. RUS HOELZEL[2], RALPH TIEDEMANN[3], JOHN W. BICKHAM[4], and members of the IWC working group.
[1]*Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Blvd. East, Seattle WA 98112, USA, robin.waples@noaa.gov.* [2]*School of Biological and Biomedical Sciences, Durham University, Durham DH1 3LE, UK, a.r.hoelzel@dur.ac.uk.* [3]*Institute of Biochemistry and Biology, University of Potsdam, D-14476 Potsdam,Germany, tiedeman@uni-potsdam.de.*
[4]*Department: Forestry and Natural Resources, Purdue University, 715 West State Street,West Lafayette, IN 47907, USA, bickham@purdue.edu.*

ABSTRACT
Recently, an IWC workgroup developed guidelines for quality control of DNA data. Once data have been collected, the next step is to analyze the data and produce results that are useful for addressing practical problems in the management of cetaceans. This is a complex exercise, as numerous analyses are possible and users have a wide range of choices of software programs for implementing the analyses. Here, we provide an outline for a document that would provide guidelines for analysis and interpretation of genetic data in a management context. We include a few worked examples to illustrate the type of content the document might contain. We encourage comments and suggestions from managers, cetacean biologists, and geneticists to help make the final document as useful as possible.

INTRODUCTION
Recently, guidelines were adopted for quality control of DNA data intended for use within the International Whaling Commission (IWC 2009). Once the data have been collected, the next step is to analyze the data and produce results that are useful for addressing practical problems in the management of cetaceans. This is a complex exercise for two major reasons: 1) a large number of methods can be used to analyze genetic data, and an equally wide range of software programs are available for conducting these analyses; and 2) a key objective is to inform those involved in cetacean management who don't have a background in population genetics. For these reasons, it has been suggested that it would be useful to have a document that provides guidelines for the analysis of population genetic data for use in a management context. Although it is not possible (nor is it desirable) to prescribe specific procedures for all analyses of population genetic data, it can be useful to provide general guidelines for some of the more common types of analyses conducted in a management context. The latter is the objective of this paper. The emphasis will be on a general discussion of issues involved in genetic data analysis rather than detailed comments about specific software programs, but some popular programs will be discussed to make particular points. Because a large number of types of analyses (and software packages to conduct such analyses) are available, to focus on analyses most relevant to a particular study we organize the discussion around some common management problems one might try to address with genetic data. These problems are identified below with roman numerals. We assume that before the analyses considered here begin, the DNA quality control guidelines (IWC 2009) have been consulted and followed to the extent possible, and that any substantial deviations have been documented and explained.

We use Bryde's whales to illustrate how genetic data might be applied to a practical management problem within the IWC framework (see Appendix 1 for more details). Bryde's whales form a taxonomic complex, with distinct offshore and coastal forms in several places. In the Pacific, separate genetically-distinct offshore population(s) with mirror-image behaviour exist in the Northern and Southern Hemispheres. Our example concerns Northern Hemisphere feeding grounds between 130°E (W Japan) and 155°W/205°E (Hawaii) and about 10°N to 40°N, where the IWC is currently considering proposals for an ongoing catch of Bryde's whales. Samples are not available from the breeding grounds (believed to be further south), and it is uncertain whether multiple stocks occur within the feeding grounds. This is a crucial issue, as the RMP quota-setting rule has been shown to have minimal risk of causing unacceptable depletion if the affected area represents a single freely-mixing population. However, if multiple stocks exist in the area and the entire quota were to be taken from a sub-area containing mostly one subpopulation, the result could easily be an unacceptable local depletion in that sub-area. The question before the Scientific Committee, therefore, is whether it is necessary to divide the catch quota spatially across this large area to minimize risk of local depletions. Three sub-areas are under consideration (#1W=130°E-165°E, #1E=165°E-180°E, #2=180°E-205°E/155°W; see map in Appendix 1).

Currently available biological information is not definitive regarding stock structure. Distribution is continuous across the catch area, and no geographically-correlated morphometric differences have been detected. However, differences in age composition have been found between the three sub-areas, and this helped motivate consideration of the possibility stock structure. Mark-recapture data documents movements between the two western subareas, but information is lacking for area #2. No genetic data are available for area #2 either. Allozyme data differentiate WNP Brydes' and other Pacific Brydes' whales but do not show differences between the two western subareas, nor do mtDNA and microsatellite data. No genetic samples are available from the breeding grounds.

Five stock structure hypotheses are currently under consideration; they differ in the number of putative stocks and the extent of hypothesized mixing between them. Potential uses of genetic data include the following:

- to evaluate the relative plausibility of the various hypotheses currently under consideration;
- to help parameterize quantities such as mixing proportions (of breeding sub-populations in feeding sub-areas) and dispersal proportions (migration between sub-populations);
- to consider whether the range of hypotheses needs alteration/expansion.

One major challenge will be to convert genetic estimates of gene flow (units = individuals/generation) into demographic estimates of more practical relevance to management (e.g., the fraction of the population that migrates between subareas each year). The technical challenges involved are discussed in detail in Sections II-V of this report.

GUIDELINES

I. Species identification/delimitation

Issues related to alpha taxonomy come up consistently, especially regarding the boundary between populations and species of small cetaceans. This section should distinguish between species identification and species delimitation, and will have some overlap with Section IV. Because a standardized methodology for DNA-based species identification of cetaceans already exists (Baker et al. 2003; Ross et al. 2003), this document will focus on analyses of genetic diversity within species. Information about *DNA Surveillance* and the comprehensive reference database, *Witness for the Whales,* can be found at the following url: http://www.cebl.auckland.ac.nz:9000/. Ross and Murugan (2006) present results of a comparison of cetacean DNA sequences in *Witness for the Whales* and *GeneBank*.

II. Analysis of diversity within populations/species

A. Information related to tests of Hardy-Weinberg (HW) equilibrium. We assume that HW evaluations have been conducted as part of the DNA data quality control step; here, attention would focus on HW deviations that might provide insight into biological processes (such as inbreeding or population mixtures).
B. Information related to tests of linkage disequilibrium (LD). As was the case for HW, the focus would be on signals that might provide insight into biological processes.
C. Measures of genetic diversity, including rarefaction (controlling for sample size in estimating allelic richness)

III. Estimating population size

A. Census size, N
    1. DNA mark-recapture
    2. Analysis of close relatives
    3. Identifying recent population bottlenecks
B. Effective population size, N_e
    1. Historical N_e (including coalescent)
    2. Contemporary N_e
        a. Single-sample methods
        b. Two-sample (temporal) methods
    3. Recent vs ancestral N_e using isolation with migration models
    4. Signals of historical population expansion/contraction
C. N_e/N ratios

IV. Analysis of diversity among populations (aka stock structure)

This is probably the most common type of management problem that utilizes genetic data. This section might be organized in any of several different ways. One way would be to use a framework based on the major stock

archetypes identified in TOSSM. The approach below uses as a first cut whether or not individuals can plausibly be grouped into putative populations before analyses are conducted.

A. Putative populations defined *a priori*.

In this case, the analyses are conducted on groups of individuals (samples); individuals are grouped into samples based on collection location, assumptions about geographic configuration of populations, or other *a priori* hypotheses about population membership.

1. Testing panmixia
2. Describing population structure
    a. F_{ST}, genetic distance, and related measures
    b. Ordination
    c. Isolation by distance/Landscape genetics (units = samples)
3. Estimating migration
    a. Methods that assume migration-drift equilibrium and estimate long-term patterns of gene flow (mN_e)
        i. Island model
        ii. Stepping-stone models
        iii. Coalescent methods and directional gene flow
    b. Assignment methods that estimate contemporary migration rate (m)
    c. Isolation with migration models to estimate splitting times and post-division migration rates
4. Mixture analysis (e.g., resolving stock composition of samples from feeding grounds or migration pathways)
5. Kinship

B. No *a priori* basis (or a questionable basis) for grouping individuals into populations

In this case, the analyses are conducted on individuals rather than groups of individuals.

1. Clustering programs
2. Clustering based on ordination
3. Landscape genetics (units = individuals), including temporal structure
4. Kinship

V. Generic/cross-cutting issues

Some issues will apply to many of the above analyses. Examples include:

A. Choice of markers (mtDNA, microsats, SNPs)
B. Ascertainment bias
C. Multiple testing
D. Mutation rates
E. Sampling and experimental design, including confidence intervals
F. Underlying assumptions and sensitivity to their violation
G. Bayesian vs maximum likelihood vs frequentist methods
H. MCMC issues (burnin, convergence)
I. Integrating genetic and non-genetic data
J. Possible influence of selection
K. Interpreting negative results
L. Validation/transparency

VI. Summary and conclusions

EXAMPLES

The following examples illustrate the type of treatment and level of detail that might be used for individual topics in the guidance document. In the next iteration of this document, we plan to relate each of these examples to the Bryde's whale example to illustrate practical application of specific types of analyses.

IV.A.2.a. F_{ST} and related measures

The so-called fixation index, F_{ST}, is one of the F-statistics originally suggested by Sewall Wright (see, e.g., Hartl & Clark 1989). It is a measure of population divergence under a scenario where one investigates whether a population is genetically subdivided into partially or completely isolated subpopulations. These putative subpopulations have to be defined *a priori*. F_{ST} then can be used to evaluate whether gene frequencies are indicative of restricted gene flow across these subpopulations (rejection of panmixia). In conservation and management, F_{ST} is hence a tool when using molecular data to identify "stocks", "management units", or "units to conserve".

The basic formula for F_{ST} is:

$$F_{ST} = \frac{H_{eT} - H_{eS}}{H_{eT}} \quad (1)$$

where H_{eT} is the expected heterozygosity in an entire population and H_{eS} is the mean expected heterozygosity within the subpopulations. The expected heterozygosity is the probability that a randomly taken individual is heterozygous at a given locus. It is calculated from allele frequencies and ranges from 0 (all individuals are expected to be homozygous, meaning that there is only a single fixed allele at an investigated locus) to 1 (all individuals are expected to be heterozygous, meaning that all alleles at an investigated locus are different). The underlying logic of F_{ST} is as follows: Under the null hypothesis of no divergence among subpopulations, H_{eS} measures for single subpopulations are unbiased estimators of the heterozygosity H_{eT} of the entire population. Hence, $H_{eT} - H_{eS} = 0$ and as a consequence F_{ST} also equals 0. Maximum divergence is reached if every subpopulation is fixed for one allele. Then, $H_e = 0$ in every subpopulation, so $F_{ST} = H_{eT} / H_{eT} = 1$.

If more than one locus is investigated, different ways of calculating multilocus F_{ST} (either averaging single-locus F_{ST} values, or averaging expected heterozygosities across loci and using those average values in Equation 1) can yield different values. In its original formulation, F_{ST} was restricted to autosomal loci of diploid organisms. However, the expected probability of a randomly taken diploid individual to be heterozygous at a given locus is conceptually identical to the probability that two randomly taken haplotypes of a locus of any ploidy level are different from one another, a measure called haplotype diversity or (more generally) gene diversity (Nei 1987).

Fixation indices calculated from haplotype and/or gene diversity are sometimes identified with a different symbol (G_{ST}), but – as formula and concept are identical – most authors have adopted F_{ST} as the "one-fits-all" name, regardless whether it is calculated on the basis of heterozygosity, haplotype diversity, or gene diversity. Because the theoretical distribution of F_{ST} remains unknown, significance of any deviation from the null hypothesis of panmixia is typically evaluated by permutation analysis. Rejection of the null hypothesis means that F_{ST} is significantly larger than 0, indicating that the subpopulations exchange fewer individuals than are expected under random dispersal (=panmixia).

The strength of F_{ST} as a measure of population divergence is threefold: (a) it is easily calculated and tested for significance for various types of molecular data; (b) the interpretation is straightforward, i.e., adoption ($F_{ST} = 0$) or rejection ($F_{ST}>0$) of panmixia; and (c) it provides a measure of the fraction of total genetic variance that is apportioned among subpopulations. This has made the F_{ST} a very commonly used measure of genetic divergence.

There are, however, a number of known limitations to F_{ST}:

1. It is based on diversity estimators (i.e., heterozygosity, haplotype diversity) that solely classify alleles/genotypes according to identity/non-identity, without any consideration of the level of difference among the alleles/genotypes.
2. The theoretical maximum of $F_{ST} = 1$ can only be reached if each subpopulation is fixed for a single unique allele. If there is variability within any subpopulation, the maximum F_{ST} is $(1 - H_{eS})$. Unfortunately, this limit to the maximum F_{ST} is often overlooked. The maximum value for F_{ST} is the smaller the more variable a marker is, and the effect can be especially dramatic for microsatellites, which often exhibit high H_{eS} (over 0.9, in which case the maximum F_{ST} is only 0.1). In the extreme (yet possible) scenario of two subpopulations completely divergent (i.e., not sharing a single allele), but both with H_{eS} approaching 1 (i.e., all individuals are expected to be heterozygous because of high allelic diversity), F_{ST} becomes meaningless, as its theoretical maximum is then 0 (see Fig. 1 in Jost 2008 for a graphical representation).
3. While the significance of F_{ST} is a good indicator of deviations from panmixia, the same information can be obtained more directly by testing equality of allele frequencies in different samples (Section IV.A.1).
4. Translation of F_{ST} into dispersal estimates, although mathematically straightforward for populations assumed to be in equilibrium between genetic drift and gene flow, is very challenging for most real-world situations (see Section IV.A.3). Given this limitation and the effect of H_{eS} on F_{ST}, the numerical value of F_{ST} should be interpreted with caution. In particular, comparing F_{ST} measures across studies will not necessarily provide a meaningful comparison of the respective level of divergence.

Alternative measures of divergence have been suggested to overcome some of the limitations of F_{ST}:

1. Some measures of population divergence incorporate the degree of divergence among genotypes. The F_{ST}-related measure Φ_{ST} explicitly accounts for mutational differences (Excoffier et al. 1992). For sequence data, divergence can also be expressed as nucleotide diversity (π) according to

$$\pi = \sum_{i,j} x_i \, x_j \, \pi_{ij,} \qquad (2)$$

where x_i and x_i are the relative frequencies of the alleles/genotypes *i* and *j* and π_{ij} is the sequence divergence among these two alleles/genotypes (Nei 1987). Analogously to H_{eT} and $\hat{H}_{eS}$, one can calculate π_T as nucleotide diversity for the entire population and mean π_S as the average nucleotide diversity within subpopulations. Subtracting mean π_S from π_T yields a measure of average % sequence difference due to divergence among subpopulations (Quinn & White 1987). For microsatellites, one can reasonably assume a step-wise mutation model, in which case alleles of similar allele size are likely to be more closely related. This is incorporated in the divergence measure R_{ST} (Slatkin 1995):

$$R_{ST} = \frac{\overline{S} - S_w}{\overline{S}} \quad , \qquad (3)$$

where S_w is the sum over all loci of twice the weighted mean of the within-population variances in allele size within subpopulations and $\overline{S}$ is the sum over all loci of twice the variance in allele size in the entire population.

2. From a mathematical point of view, the theoretical range of possible F_{ST} values can be easily re-adjusted to the interval [0,1] by a correction factor (Hedrick 2005):

$$F_{ST\,adjusted} = \frac{H_{eT} - H_{eS}}{H_{eT}(1 - H_{eS})} \qquad (4)$$

This correction, however, has not been widely applied. In addition, there is so far neither a theoretical nor an empirical evaluation of the impact of such a correction on the distribution of F_{ST}. As a consequence, it remains unclear whether $F_{ST\,\text{adjusted}}$ also suffers from the limitations mentioned under (3)—specifically, that identical values do not necessarily imply an equal level of divergence.

3. It has been recently argued that a true measure of differentiation should decompose overall diversity into two parts: (a) the variance within the subpopulations and (b) the variance because of divergence among subpopulations (Jost 2008). To yield an unbiased measure of the divergence component in heterozygosity, this author suggests a formula very similar to (IV), i.e.,

$$H_{ST} = \frac{H_{eT} - H_{eS}}{1 - H_{eS}} \, , \qquad (5)$$

where H_{ST} is the heterozygosity component caused by divergence among subpopulations. To yield the final estimator of differentiation (*D*) with values on the interval [0,1], H_{ST} is multiplied by (n/(n-1)) where n is the number of subpopulations (Jost 2008). As neither H_{ST} nor *D* are implemented in the most widely used data analysis software packages, these measures are rarely used. In addition, Jost's *D* has been found (a) to be (like F_{ST}) also affected by levels of heterozygosity, and (b) not to be conceptually related to basic population genetics quantities (effective population size, gene flow), rendering it potentially inferior to F_{ST} as a standard measure of population differentiation (Ryman & Leimar 2009).

In summary, F_{ST} is an easily calculated measure of population differentiation and a good indicator of the presence/absence of panmixia. It has, however, limitations due to (a) neglect of the amount of genotype divergence and (b) dependence of the divergence estimator upon the within-subpopulation diversity. Alternatives to F_{ST}, which could at least partially overcome these limitations are available, but some are only rarely used and/or not generally accepted (in particular, H_{ST} and *D*). Given its widespread application and its sound theoretical foundation, F_{ST} can be considered as a valuable measure to infer population structure from genetic data, if *a priori* information about putative subpopulations is available. However, it is recommended to bear the limitations of F_{ST} in mind and to treat numerical comparisons among F_{ST} values with caution.

IV.A.3.b. **Assignment methods that estimate contemporary migration rate (*m*)**

Efforts to estimate levels of connectivity from genetic data have traditionally relied on equilibrium models that integrate information over evolutionary time periods (see section IV.A.3.a.). The last decade has seen increasing interest in so-called 'assignment methods' that do not require equilibrium assumptions and instead can estimate

contemporary patterns of migration over time frames encompassed by the samples. 'Assignment tests' (Paetkau 1995; Manel et al. 2005) are a type of discriminant function analysis in which the discriminant functions are based on genetic traits that differ in frequency among potential source populations. Multilocus genotypes are used to 'assign' individuals to the most likely source population, guided by learning samples collected from potential sources. If an individual is assigned to a population other than the one it was sampled from, it can be inferred that the individual is a migrant (Waser and Strobeck 1998; Berry 2004). The program *GeneClass* (Piry et al. 2004) includes several different assignment test methods and offers the user various options for attempting to identify first-generation migrants. Other programs attempt to identify second-generation migrants (Wilson and Rannala 2003) or estimate the fraction of genes in each individual that are derived from each population (Pritchard et al. 2000).

Assignment methods have some advantages for estimating migration: they don't require one to assume migration-drift equilibrium, as do most standard models; they can potentially provide very detailed information about connectivity (both magnitude and direction); and they provide information about contemporary dispersal, which might be of interest for a variety of reasons. However, assignment methods also have some substantial limitations for studying dispersal. First, these methods provide information about movement of individuals but not reproductive success of the migrants; therefore, they do not provide any direct information about gene flow. Second, assignment methods provide information about dispersal only for the time frames encompassed by the sampling. Because dispersal is a stochastic process, samples taken from only one or a few years might not provide a representative picture of migration. This can be contrasted with equilibrium models, which can provide an estimate of long-term patterns of gene flow from samples taken at a single point in time. For any given application, these two factors might or might not represent serious limitations, depending on the nature and objectives of the research program.

A third factor—statistical power—is potentially a more general limitation on use of assignment methods to study contemporary dispersal. Power to detect migrants with genetic methods depends on two things: the amount of data one has (samples of individuals, gene loci, and alleles), and the magnitude of genetic differences among populations. The researcher has control over the former but not the latter, and therein lies a conundrum: power is highest when genetic differences among populations are large, but in that case migrants will be rare and difficult to detect without a very ambitious sampling program; conversely, if migration is high enough to provide reasonable prospects for finding migrants, the resulting levels of gene flow should erode most differences among populations, making it difficult to genetically distinguish migrants from residents.

Two examples illustrate the inherent difficulty related to power. Paetkau et al. (2004) used computer simulations to evaluate power to detect first generation migrants. They found that even with fairly large amounts of data (50 individuals sampled per population; 20 microsatellite-like gene loci), power to detect true migrants was <50% when gene flow rates were high enough ($mN \geq 5$) to keep F_{ST} values below about 0.05. These conditions would apply to a substantial fraction of potential applications for cetaceans.

Second, the power issue sets up an inherent tradeoff between Type I (incorrectly labeling a resident as a migrant) and Type II (failing to detect a true migrant) error rates, either of which can seriously bias estimates of migration. Consider this hypothetical example: a group of populations with $N = 100$ individuals each are connected by 1% migration per generation ($m = 0.01$). This leads to $mN = 1$ (a low level of gene flow) and relatively large genetic differences among populations. Optimistically, assume that these large differences lead to ~100% power to detect migrants using assignment methods (as found by Paetkau et al. 2004 for data-rich scenarios). So, a large sample would on average contain 1% true and correctly-identified migrants. But if the standard tolerance for Type I error is used ($\alpha = 0.05$), then 5% of the sample would also be incorrectly identified as migrants. In this case, even with perfect statistical power, the estimate of migration rate (0.01 + 0.05 = 0.06) would be six times the true level. The only solution to this problem is to adopt a very low α level, but doing so is likely to compromise power unless genetic differentiation is very strong.

The conundrum regarding power does not necessarily represent an insurmountable problem for using assignment methods to study contemporary dispersal—for example, Berry et al. (2004) reported reasonably good agreement between genetic and mark-recapture estimates of dispersal in a series of populations of the grand skink, *Oligosoma grande*, for which F_{ST} values ranged between 0.04 and 0.11. However, the issues discussed above do indicate that careful attention to experimental design is essential, as is a realistic assessment of prospects of producing useful information. Three general strategies can help improve performance. First, in theory at least, it is possible to achieve high power to identify migrants among populations with very low levels of genetic differentiation, provided

that arbitrarily large numbers of loci and alleles can be scored. At present the ability to do this with non-model species is limited, but that situation could change in the future. Second, adopting a very low tolerance for Type I errors (e.g., $\alpha \leq 0.01$) can help reduce some of the most serious sources of potential bias, but this will likely compromise power unless genetic differences are moderately large and/or very large amounts of data are available. Third, the major challenges to these methods arise from uncertainty in identifying individual migrants. Using an analogue to Genetic Stock Identification (resolution of mixed-stock fisheries using genetic data—Shaklee et al. 1999; Manel et al. 2005), if focus is shifted from identifying individual migrants to estimating an overall migration *rate*, then uncertainty about origins of individuals might not preclude precise and accurate estimates of migration. This would require developing, or at least refining, some new analytical techniques. One software program, *BayesAss* (Wilson and Rannala 2003) does actually attempt to estimate migration rate, but its performance with weakly differentiated populations has not been encouraging (Faubet et al. 2007).

Finally, the conundrum regarding the inverse relationship between the level of migration and genetic differentiation largely disappears if the system one is analyzing involves populations that historically have been strongly isolated (and hence are well differentiated genetically) but which are currently exchanging sizeable numbers of migrants. This non-equilibrium situation cannot persist for long unless the migrants have little or no reproductive success, but in the interim could provide a large number of migrants to sample *and* high power to distinguish them from residents. This scenario, in fact, is one that the Wilson and Rannala (2003) program was designed to study. It is not clear how often this scenario might occur with cetaceans.

IV.A.3.c. Isolation with migration models

Evolutionary models that form the basis for genetic comparisons of different populations have typically been based on either of two extreme scenarios. One model considers populations that have been exchanging migrants at a constant rate for an effectively infinite period of time (an equilibrium migration model). The other considers populations that descended from some common ancestral population at some point in the past, and have since then evolved independently without gene flow (an isolation model). Many of the population comparisons in the literature use measures based on Wright's inbreeding coefficients (especially F_{ST}; Wright 1965—see Section IV.A.2.a). The problem is that, by itself, a measure of genetic distance such as F_{ST} can't differentiate between the two scenarios described above. A low F_{ST} could mean either a relatively high rate of gene flow over time, or the recent cessation of gene flow altogether (because differentiation by genetic drift or selection will take some time to accumulate). Isolation with migration (IM) models allow both cases to be considered together and therefore distinguished. This is important from a management perspective when trying to distinguish population segments for which there is some ongoing connectivity, from populations that have become isolated (and are potentially incipient species). Boundaries to gene flow can be difficult to identify in the ocean, and there have been several examples of cryptic cetacean species being identified in recent years (e.g. Wada et al. 2003). There are also a number of poorly resolved radiations of species that are polyphyletic (e.g. the *Tursiops, Delphinus, Stenella* complex), and nominal species that have been variously recognized as single species or divided into multiple sub-species or putative species (e.g. *Delphinus delphis* & *Orcinus orca*; see below for an example). IM can help resolve these questions associated with alpha taxonomy, and thereby better identify management units. IM models can also help identify the relevant mechanisms and processes that regulate gene flow, and thereby improve the efficiency of management strategies.

An early example of an IM model was described by Wakeley (1996), based on the observation that the variance of pairwise nucleotide differences (a measure of the extent of diversity) is smaller under isolation than under equilibrium migration. However, any single statistic such as this will necessarily leave out much of the complexity likely to be reflected in real demographic histories. This fact inspired an approach that could simultaneously estimate six parameters, initially described by Nielsen & Wakeley (2001). It is a two-population, one-ancestor model, and the parameters estimated are the effective population size of each of the three populations (N_1, N_2 & N_A), the migration rate from population 2 into population 1 (m_1), the migration rate from population 1 into population 2 (m_2), and the splitting time between the two extant populations (t). Thus, the IM model simultaneously estimates parameters related to both the equilibrium-migration and isolation models. The input is sequence data from a single locus, and the model assumes an infinite allele model of mutation (all new mutations are novel), selective neutrality, constant population size, no recombination, and a closed system (no other populations exchanging alleles with the focal populations). The model was tested in both likelihood and Bayesian frameworks. A later version allowed for the application of a finite site model of evolution (Palsboll et al. 2004). This was useful, as this approach is especially appropriate for the analysis of mtDNA sequence data (no appreciable recombination), but a finite site

mutation model (such as HKY; Hasegawa et al. 1985) is better suited to mtDNA where mutation rates are comparatively high. Another limitation of the original model was addressed in a paper by Hey & Nielsen (2004). The Nielsen & Wakely (2001) method (implemented in the program MDIV) was based on a single locus, but single gene trees can be affected by stochastic processes or natural selection and become unrepresentative of the true history.

Hey & Nielsen (2004) extended the IM method to include multiple loci (implemented in the program IM). The parameterization is also different in this model, where population size, migration rate and splitting time are all scaled by mutation rate. It is assumed that the loci included have been drawn from all loci at random (that is, that none are atypical with respect to the depth of the gene tree or degree of gene flow), and as before, that they have evolved neutrally. The basic parameters measured are the same as before, but additional parameters are required to accommodate the additional loci (scalars to account for differences in mutation rate and the mode of inheritance). As described in a later paper (Hey 2005), the IM program can also model populations that are expanding or contracting (through the addition of the 's' parameter). Three mutation models are available in the IM program: the infinite sites model (often appropriate for nuclear genes that evolve relatively slowly), the HKY finite-site model (better suited for relatively fast evolving sequences, such as mtDNA), and the stepwise mutation model (appropriate for microsatellite DNA loci). Note that mutation rates are input as per locus per year.

The programs that test these models employ Markov chain Monte Carlo (MCMC) simulations, and a major challenge with this approach is to verify that the output corresponds to something like reality. It is, unfortunately, quite easy to generate data that looks plausible but is in fact quite wrong. Therefore, a number of precautions are necessary, including the need to run the same simulations at least three times. The first time allows some assessment of the effectiveness of the chosen input parameters, while the last two full runs allow confirmation that independent runs (with different random number seeds) give equivalent results. While a detailed explanation is beyond the scope of this summary, the key objective is for the simulation to converge on the 'stationary distribution' (the distribution that you want to estimate with your sample of parameter values). This convergence is necessary if the results are to be credible, and therefore a variety of indicators need to be assessed during the progress of the run. A description of these indicators and a more detailed explanation of the problem can be found in the support documents for the program IM. A consequence of this is that the simulations sometimes need to run for a very long time (often weeks or more, depending on the speed of the computer processors). One feature in the program IM that helps is the ability to run a number of chains in parallel (called 'Metropolis coupling'). A further advance was implemented in a second version of the program (IMa; Hey & Nielsen 2007), whereby approximations of the values for some parameters allows the program to progress more quickly.

A few applications have been published for cetacean species. In one study fin whales inside and outside the Mediterranean basin were compared, and a model of ongoing gene flow (at about 2 females per generation) was shown to be better supported that the alternative of recent isolation (Palsboll et al. 2004). In another study the timing of the founding of populations of killer whales currently using coastal habitat was found to post-date the last glacial maximum (after habitat was released from under ice), and ongoing gene flow was indicated between populations of different ecotypes in the eastern North Pacific (Hoelzel et al. 2007). These same ecotypes have recently been proposed as different species (Morin et al. 2010), though this was based only on mtDNA. Nuclear markers suggested ongoing gene flow both from IM (Hoelzel et al. 2007) and individual genotype and parentage analyses (Pilot et al. 2010). An important limitation remains with the model as applied in the program IMa. The two population, one ancestor model means that there should not be other populations more closely related to the sampled populations than they are to each other, and that there shouldn't be unsampled ('ghost') populations exchanging genes with either the focal populations or the ancestral population. Violations of these assumptions can inflate the apparent size of the ancestor population, and make estimates of gene flow misleading (which might in fact be signals of gene flow through intermediaries). To help address this problem, a new version of IMa (IMa2) has now been released which allows inclusion of up to 10 populations (including ghost populations; Hey pers comm). The difficulty with using a multipopulation model, however, will be the introduction of new parameters to estimate, as well as more expansive requirements for sampling. A modification of this type would need to be supported by a large number of highly informative loci to produce reliable results; even then, obtaining reliable estimates for closely related populations is likely to be very challenging.

IV. B. 3. Landscape genetics

Landscape genetics merges the fields of molecular population genetics, landscape ecology and spatial statistics to study how features of the landscape affect processes of microevolution such as gene flow, natural selection, local adaptation and genetic drift (Manel et al., 2003; Storfer et al. 2007; Holderegger and Wagner 2008; Segelbacher et al. 2010). Unlike most traditional approaches to the study of population subdivision, landscape genetics does not necessarily require the *a priori* assignment of individuals to populations, although that can be done as well. Often in landscape genetics the individual is the operational unit, and thus the experimental design can be free of assumptions (and their potential biases) about population structure. This also allows for studies on a much finer scale than is possible for analyses that require individuals to be grouped together into putative populations. Landscape genetics typically focuses on contemporary evolutionary processes; in this respect, it differs from phylogeography (another approach that links genetics to geography), which seeks to identify phylogenetic lineages and thus looks at the consequences of population structure over considerable periods of time.

The general aim of landscape genetics is to identify genetic discontinuities (i.e., population structure) that correlate with landscape and/or environmental features. From a conservation biology perspective, understanding the functional connectivity of populations across landscapes is an important goal (Van Dyck and Baguette 2005). One primary application is to aid in the identification of management units, and the method can be particularly useful in cases where it is difficult to assign individuals to populations *a priori.* This is especially important for continuously distributed species, for which any *a priori* grouping into putative populations risks being arbitrary. Spatial patterns that potentially can be identified include clines, genetic discontinuities, metapopulations, isolation by distance, and random patterns. In the final analysis, validation of patterns of functional connectivity is ultimately based on gene flow estimates (Cushman et al., 2009). Obviously, obtaining accurate locality data for each individual is a critical requirement of sampling. Genetic and statistical methods are used to identify spatial genetic patterns and their correlations with landscape and environmental features. Landscape genetic analyses can use any of the typical molecular markers, including mtDNA, microsatellites, SNPs, AFLPs, and allozymes. Neutral markers are useful for identifying some ecological and demographic processes, but targeted genes undergoing selection can potentially provide novel information about the genetic basis of adaptation, adaptive differentiation, and speciation. Sampling across ecological gradients can help to identify environmental factors that might drive selection.

Landscape genetic studies in which individuals can be *a priori* assigned to populations can employ standard population genetic methods such as F-statistics (see Section IIV.A.2.a) and assignment tests (see Section IV.A.3.b). However, in cases where the individual is the operational unit, application of Wright's Neighborhood concept can be useful, particularly for detecting patterns of isolation by distance. For continuously distributed species, Wright (1943) defined the genetic 'neighborhood' as the basic unit of population structure. In Wright's model, neighborhood size is $4\pi D\sigma^2$, where D is population density and σ^2 is mean squared parent–offspring distance along one axis. Statistical procedures for use in landscape genetics include Mantel's test for isolation by distance. Regression analysis of genetic distance with geographic distance allows the estimation of neighborhood size and thus dispersal distance. Spatial autocorrelation analysis allows one to identify spatial patterns such as clines and test whether distance is the main determiner of population structure. Bayesian clustering methods (see Section IV.B.1) can be used to assign individuals to populations by a method that minimizes HW or linkage disequilibria among groups. Multivariate analyses like Principal Components Analysis (PCA) are used to summarize variation among many loci across an area. Interpolation of the major components into a synthesis map gives insight into spatial patterns and can allow identification of patterns such as clines. Two popular methods (Barbujani, 2000) to identify genetic boundaries or discontinuities are Monmonier's algorithm, which visualizes data contained in a genetic distance matrix on a geographical map to identify boundaries, and wombling, which "locates boundaries across a surface for an interpolated variable (i.e. allele frequency surface) by searching for regions in which the absolute value of the surface slope is large" (Manel et al., 2003).

Studies of landscape genetics often apply methods of Bayesian clustering to more objectively group individuals into populations. The program STRUCTURE (Pritchard et al. 2000) is the most widely applied Bayesian clustering approach. However, because of limitations to this program and to the use of isolation by distance models in general, Segelbacher et al. (2010) suggest that a new trajectory for investigations in landscape geneticists will be the study of population genetic structure and isolation by distance under computer-simulated dynamical models.

Statistical approaches also exist that allow for the correlation of genetic patterns with environmental variables. Mantel tests are used to correlate genetic distance to a wide variety of variables. Alternatively, canonical

correspondence analysis (CCA) can be used to relate genetic diversity to environmental factors, at the same time testing for environmental factors that explain variation in genetic diversity. Geographical information systems (GIS) visualize spatial genetic patterns. Because GIS allows overlay of environmental or landscape variables onto genetics data, it facilitates development of hypotheses about the cause of spatial genetic patterns.

There are limitations and constraints to the use of landscape genetics, as with all analytical methods. Some analyses assume random mating, which might be violated by many cetacean populations. In addition, gametic disequilibrium and departures from HW proportions can result from processes besides population structure, including small populations, bottlenecks, inbreeding and admixture; hence, these factors can obscure patterns related to population structure. Schwartz and McKelvey (2009) showed that the program STRUCTURE identified varying numbers of clusters which were dependant on sampling scheme. STRUCTURE may thus provide misleading population assessments because genetic gradients or similar patterns likely are common in nature. A more general limitation is that isolation-by-distance mathematical models strongly depend on the assumption of migration-drift equilibrium (Rousset 1997). Studies of great whales will often deal with non-equilibrium populations. Indeed, as a final caveat, the historic overharvest of great whale populations and its resulting bottlenecks could represent a significant violation of the assumptions of certain analyses, particularly those that assume demographic and genetic equilibrium.

V.C. Multiple testing

Some analyses routinely involve multiple tests of the same hypothesis (e.g., tests of HWE and LD, or pairwise tests of heterogeneity between populations). In these applications, it is common practice to use a correction for multiple testing, such as the Bonferroni correction, in which the critical *P* value is inversely proportional to the number of tests. Two points should be kept in mind when using this type of correction for multiple tests.

1) The Bonferroni correction is widely known to be conservative and hence will fail to detect some actual departures from the null hypothesis.
2) If the correction is performed, then the expectation is (with probability 1-*P*) that the number of adjusted significant tests will be zero. Therefore, even a single adjusted significant test cannot easily be attributed to chance and requires an explanation.

If a multiple testing correction is performed, a better option might be the false discovery rate (FDR; the fraction of tests in which the null hypothesis is falsely rejected; Benjamini and Hochberg 1995), which adjusts for multiple testing without sacrificing as much power as the Bonferroni correction. In addition, it is recommended that results are also presented for unadjusted tests, as the distribution of unadjusted *P* values provides valuable information about agreement with the underlying null hypothesis.

V.D. Mutation rates

The parameter $\theta = 4N_e\mu$ plays a key role in both theoretical and applied population genetics. θ is a composite parameter, proportional to the product of effective population size (N_e) and mutation rate (μ). Although this fact adds complexity to some analyses, it can be used to advantage by a simple rearrangement of the above equation:

$$N_e = \theta/(4\mu). \quad (6)$$

This means that if θ can be estimated from population genetic data (as is routinely done with microsatellites or sequence data for mitochondrial or nuclear DNA), then insights into N_e can be obtained if one can also estimate mutation rate. The effective size that is estimated in this way is a long-term, or historic, N_e that depends (among other things) on the assumption that measured levels of genetic diversity reflect an equilibrium between mutation and genetic drift (see Section III.B.1. above). This general approach has a variety of practical applications, such as estimating historical effective population size; estimating divergence times between populations or species; and estimating population demographic patterns over time (see Beaumont and Rannala 2004, Nielsen and Beaumont 2009). Several factors, however, contribute to uncertainty and limit the practical usefulness of these approaches. First, only four kinds of DNA bases occur (termed A, T, C, G for short), and DNA sequences are typically compared by counting the fraction of sites at which they have different bases. Once a mutation has occurred at a particular site (e.g., from A to T), a subsequent mutation at that site will still result in only a single difference compared to the reference sequence (if the mutation is from T to G or C) or will negate the original change (if the mutation is a back mutation from T back to A). This saturation effect is of particular relevance for estimates of mutation rate at mtDNA, which are typically obtained by the 'phylogenetic method' that involves comparing sequences from different species. In addition to making duplicate mutations more likely, this introduces potential sources of error in developing calibration points for divergence times—typically derived from the fossil record, which is relatively poor for cetaceans.

Second, mutation rates can vary considerably among species and among regions of the genome within species. For many years, a '2% rule' was used for mtDNA, based on calculations using the phylogenetic method, suggesting that for vertebrates the average rate of base substitution was about 2% per million years (Wilson et al. 1985). However, rates vary among regions of the mtDNA molecule, and for the widely used control region, the estimates were considerably higher (12% to 38% per million years in humans; see review in Henn et al. 2009). Furthermore, the mtDNA control region itself is heterogeneous for mutation rate, with the central, very conserved, region being flanked by two 'hypervariable' regions (HVR1 and HVR2).

Finally, recent estimates of mutation rate over shorter time frames for intraspecific comparisons often differ greatly from those based on the phylogenetic method. An extensive analysis (Howell et al. 2003) provided an estimate for the human HVR1 of 95% per million years (0.95 changes / site/ million years). Similar approaches applied to other species including *C. elegans* (Denver et al. 2000) and Adélie penguins, *Pygoscelis adeliae* (Millar et al. 2008) have also produced estimates that are 1-2 orders of magnitude higher than suggested by the 2% rule. Henn et al. (2009) suggested that for humans the high mutation rates decay after about 15,000 years, but for penguins the elevated rate seemed to extend back further in time (Millar et al. 2008).

From the form of Equation 6, it is easy to see that errors in estimating the mutation rate directly translate into the same proportional errors in estimating N_e. For example, if mutation rate is underestimated by a factor of two, N_e will be overestimated by the same amount. This fact, together with the wide range of published estimates of mutation rates, has helped spawn much of the controversy that has surrounded some attempts to estimate historical N_e based on existing levels of genetic diversity. For example, Roman and Palumbi (2003) calculated that there must have been many more whales in pre-whaling oceans than had previously been thought, based on an estimate of mutation rate in the cetacean mtDNA control region derived from the phylogenetic method (e.g., Hoelzel et al. 1991)—about 2% per million years. However, the relevant timeframe suggests that the much higher rate estimates derived from intra-specific genealogies might be more appropriate (e.g. Henn et al. 2009, Millar et al. 2008). If those higher rates were instead applied, the cetacean population size estimates would fall in line with what had been previously interpreted from historical catch data.

In summary, current levels of genetic diversity and other patterns in the DNA of contemporary populations contain information about historic size and demographic processes. However, deciphering this information is tricky and depends heavily on obtaining a reliable estimate of mutation rate. It is not enough to have an estimate of mutation rate for the focal species; it is also important to have estimates for the regions of the genome that produced the genetic data being analyzed, and to apply the correct mutation rate to the relevant time frame – higher rates for more recent events (Ho et al. 2005). Because of the saturation effect, estimation of mutation rates might best be confined to the time period during which each mutation occurs at a unique site and the cumulative number of substitutions increases linearly with time.

ACKNOWLEDGMENTS

Members of the IWC Workgroup on Guidelines for Population Genetic Data Analysis include Scott Baker, Mark Bravington, Greg Donovan, Mike Double, Rus Hoelzel, Jennifer Jackson, Phil Morin, Ada Natoli, Per Palsböll, and Robin Waples. Andre Punt and Mark Bravington provided the Bryde's whale example that appears in the Appendix.

REFERENCES

Baker, C.S., Dalebout, M.L., Lavery, S., and Ross, H.A. 2003. www.DNA-surveillance: applied molecular taxonomy for species conservation and discovery. *Trends Ecol. Evol.* 18:271-272.

Barbujani, G. 2000. Geographic patterns: how to identify them and why. *Hum. Biol.* 72: 133–153

Beaumont, M.A., and Rannala, B. 2004. The Bayesian revolution in genetics. *Nat. Rev. Genetics* 5: 251-261.

Benjamini, Y, and Hochberg, Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Royal Stat. Soc., Series B*, 57:289–300.

Berry O, Tocher MD, and Sarre SD. 2004. Can assignment tests measure dispersal? *Mol. Ecol.* 13, 551–561.

Cushman SA, McKelvey KS, Schwartz MK. 2009. Use of empirically derived source-destination models to map regional conservation corridors. *Conserv. Biol.* 23: 368–376

Denver DR, Morris, K, Lynch M, Vassilieva LL and Thomas K. 2000. High direct estimate of the mutation rate in the mitochondrial genome of *Caenorhabditis elegans. Science* 289:2342-2344.

Excoffier L, Smouse PE, and Quattro JM. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.
Faubet F, Waples RS, and Gaggiotti OE. 2007. Evaluating the performance of a multilocus Bayesian method for the estimation of migration rates. *Mol. Ecol.* 16:1149–1166.
Hartl DL, and Clark AG. 1989. Principles of population genetics. Sinauer, Sunderland.
Hasegawa, M., Kishino, H. & Yano, T. 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA, *J. Mol. Evol.* 22:160–174
Hedrick PW. 2005. Genetics of populations, 3rd edition. Jones and Bartlett, Sudbury, MA.
Henn BM, Gignoux CR, Feldman MW and Mountain JL. 2009. Characterizing the time dependency of human mitochondrial DNA mutation rate estimates. *Mol. Biol. Evol*. 26:217-230.
Hey, J. 2005. On the number of new world founders: a population genetic portrait of the peopling of the Americas. *PLoS Biol* 3:e193.
Hey, J. & Nielsen, R. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167:747-760.
Hey, J. & Nielsen, R. 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *PNAS* 104:2785–2790.
Ho SYW, Phillips MJ, Cooper A, and Drummond AJ. 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol.* 22: 1561–1568.
Hoelzel AR, Hancock JM and Dover GA. 1991. Evolution of the cetacean mitochondrial D-loop region. *Mol. Biol. Evol.* 8:475-493.
Hoelzel, A.R., Hey, J., Dahlheim, M.E., Nicholson, C., Burkanov, V. & Black, N. 2007. Evolution of Population Structure in a Highly Social Top Predator, the Killer Whale. *Mol. Biol. Evol.* **24**:1407–1415.
Holderegger R, Wagner HH (2008) Landscape genetics. *Bioscience* 58: 199–207
Howell N, Smejkal CB, Mackey DA, Chinnery PF, Turnbull DM and Herrnstadt C. 2003. The pedigree rate of sequence divergence in the human mitochondrial genome: there is a difference between phylogenetic and pedigree rates. *Am. J. Hum. Genet.* 72:659-670.
IWC. 2009. Report of the Scientific Committee. Annex I. Report of the Working Group on Stock Definition. Appendix 2. Guidelines for DNA data quality control for genetic studies relevant to IWC management advice. *J. Cetacean Res. Manage. (Suppl.)* 11: 252-256.
Jost L. 2008. G_{ST} and its relatives do not measure differentiation. *Mol. Ecol.* 17: 4015-4026.
Manel S, Gaggiotti O, and Waples RS. 2005. Assignment methods: matching biological questions with appropriate techniques. *Trends Ecol. Evol.* 20:136-142.
Manel S, Schwartz, MK, Luikart, G, Taberlet P. 2003. Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol. Evol.* 18: 189-197.
Millar CD, Dodd A, Anderson J, Gibb GC, Ritchie PA, Baroni C, Woodhams MD, Hendy MD, and Lambert DM. 2008. Mutation and evolutionary rates in Adélie Penguins from the Antarctic. *PLoS Genetics* 4: e1000209.
Morin, P.A., et al. 2010. Complete mitochondrial genome phylogeographic analysis of killer whales (*Orcinus orca*) indicates multiple species. Genome Research doi:10.1101/gr.102954.109
Narum SR. 2006. Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conserv. Gen.* 7:783–787.
Nei M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.
Nielsen R, and Beaumont MA. 2009. Statistical inferences in phylogeography. *Mol. Ecol.* 18, 1034–1047.
Nielsen, R., & Wakeley, J. 2001. Distinguishing migration from isolation. A Markov chain Monte Carlo approach. *Genetics* 158:885-96.
Paetkau D, Calvert W, Stirling I, and Strobeck C. 1995. Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.* 4:347–354.
Paetkau D, Slade R, Burden M, Estoup A. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Mol. Ecol.* 13:55–65.
Palsbøll, P.J., Bérubé, M., Aguilar, A., Notarbartolo di Sciara, G. & Nielsen, R. 2004. Discerning between re-current gene flow and recent divergence under a finite-site mutation model applied to North Atlantic and Mediterranean Sea fin whale (*Balaenoptera physalus*) populations. *Evolution* 58:670-675
Pilot, M., Dahlheim, M.E. & Hoelzel, A.R. 2010. Social cohesion among kin, gene flow without dispersal and the evolution of population genetic structure in the killer whale (*Orcinus orca*). *J. Evol. Biol.* 23:20-31

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, and Estoup A. 2004. *GENECLASS2*: a software for genetic assignment and firstgeneration migrant detection. *J. Heredity* 95:536–539.

Pritchard JK, Stephens P, and Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.

Quinn TW, and White BN. 1987. Analysis of DNA sequence variation. In: Avian genetics (F. Cooke & P.A. Buckley, eds). Academic Press, London, p. 163-198.

Roman J, and Palumbi SR. 2003. Whales before whaling in the North Atlantic. *Science* 301:508-510.

Ross HA, Lento GM, Dalebout ML, Goode M, Ewing G, McLaren P, Rodrigo AG, Lavery S, and Baker CS. 2003. DNA surveillance: web-based molecular identification of whales, dolphins and porpoises. *J. Heredity* 94, 111-114.

Ross HA, and Murugan S. 2006. Using phylogenetic analyses and reference datasets to validate the species identities of cetacean sequences in GenBank. *Mol. Phylog. Evol.* 40:866-871.

Ryman N, and Leimar O. 2009. G_{ST} is still a useful measure of genetic differentiation – a comment on Jost's D. *Mol. Ecol.* 18: 2084-287.

Segelbacher G, Cushman SA, Epperson BK, Fortin MJ, Francois O, Hardy O, Holderegger R, Taberlet P, Waits LP, Manel S. 2010. Applications of landscape genetics in conservation biology: concepts and challenges. *Conserv Genet* DOI 10.1007/s10592-009-0044-5.

Shaklee JB, Beacham TD, Seeb L, and White BA. 1999. Managing fisheries using genetic data: case studies from four species of Pacific Salmon. *Fish. Res.* 43, 45–78.

Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457-462.

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP. 2007. Putting the 'landscape' in landscape genetics. Heredity 98: 128–142.

Van Dyck H, Baguette M. 2005. Dispersal behaviour in fragmented landscapes: routine or special movements? *Basic Appl Ecol* 6: 535–545

Wada, S., Oishi, M., and Yamada, T.K. 2003. A Newly Discovered Species of Living Baleen Whale. *Nature* 426: 278-281

Wakeley, J. 1996. Distinguishing migration from isolation using the variance of pairwise differences. *Theor. Pop. Biol.* 49:369-386.

Waser PM, and Strobeck C. 1998. Genetic signatures of interpopulation dispersal. *Trends Ecol. Evol*. 13:43–44.

Wilson AC et al. 1985. Mitochondrial-DNA and 2 perspectives on evolutionary genetics. *Biol. J. Linn Soc.* 26:375-400.

Wilson GA, and Rannala B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics* 163:1177–1191.

Wright S. 1943. Isolation by distance. Genetics 28: 114–138.

Wright, S. 1965. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19:395-420.

Appendix 1

[Note: this is reproduced from Appendix 3 in Annex I (Report of the Working Group on Stock Definition) to the 2009 Report of the Scientific Committee]

WNP Bryde's whales: a summary for use in the Analysis Guidelines document

[More detail on the points below can be found in JCRM 9 (2007) ~p410, IWC 58 Annex D ~p95, JCRM 10 (2008) p 455]

Bryde's whales feed during the Northern summer in mid-latitudes across the North Pacific ocean, eastwards from at least 130°E. They migrate there from breeding ground(s) further south; the location of these ground(s) is unknown. Separate genetically-distinct population(s) with mirror-image behaviour exist in the Southern Hemisphere. Taxonomically, Bryde's whales form a complex, with for example distinct and small coastal forms in several places, but all the discussion here concerns the offshore form only. There may or may not be stock substructure within the feeding grounds, and/or on the breeding grounds. The IWC is considering proposals for an ongoing catch of Bryde's whales on the feeding grounds between 130°E (W Japan) and 155°W/205°E (Hawaii) and about 10°N to 40°N.

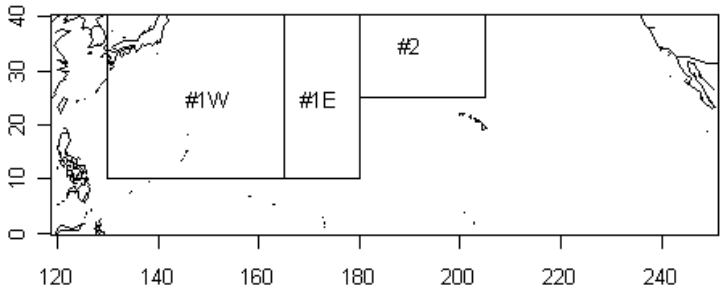

Figure 1. Areas of the North Pacific where the IWC is considering catches of Bryde's whales. See main text for meaning of boxes.

Some data on basic biological and population dynamic parameters is available.

- Bryde's whales mature around age 7, and median age is around 16.
- The abundance for areas #1W, #1E and #2 combined was estimated to be about 22000 in 1995. No reliable estimates are available outside these areas, so if these areas only constitute part of the feeding grounds of the underlying stocks, the population abundance will be higher.
- Catches have been taken since at least 1906, primarily from #1E and #1W. Annual takes were in the low hundreds until the 1970s, with roughly 1000 animals per year taken until 1987 when catching stopped. Catches of around 50 per year resumed in 2000.

In terms of non-genetic data that might inform stock structure: there is no discontinuity in distribution within the catch area, nor any evidence of morphometric differences. Differences in age composition have been found between the three sub-areas; although interpretation is unclear, this provides part of the motivation for considering the possibility stock structure. Mark-recapture shows that movements do occur across the two western sub-areas; there has not been enough effort in the easternmost sub-area, #2, to tell how much movement there is to/from #2. Two satellite tags have now been deployed, for short durations.

In terms of genetic data:

- single-locus allozyme data has shown differences between WNP Brydes' and other Brydes' whales in the Pacific, but not within the catch zones;
- mitochondrial and microsatellite DNA has been collected from the two western sub-areas in the 1980s and 2000s, with no significant heterogeneity revealed;
- no genetic samples have been collected from the easternmost sub-area;
- no genetic samples are available from the breeding ground(s), and there is little prospect of getting any.

In terms of proposed future catches: catch levels over time would be regulated as usual under the RMP, based on time series of abundance estimates collected across the catch area. The quota-setting rule has been shown to have minimal risk of causing unacceptable depletion if applied to a single freely-mixing population. However, if there was population structure on the catch grounds, and if the entire quota was taken from a sub-area containing mostly one subpopulation, then it would be possible to cause unacceptable local depletion in that sub-area over, say, a 100-year timeframe; any refill from subpopulations that primarily feed in other sub-areas would be too slow in management terms. The question before the Scientific Committee, is whether it is necessary to divide the catch quota spatially across this large area, to avoid risk of local depletions. Currently, three sub-areas are under consideration (#1W=130°E-165°E, #1E=165°E-180°E, #2=180°E-205°E/155°W).

The boundaries of these sub-areas are fairly arbitrary, biologically speaking, and they simply represent a tractable way in which catches could be divided if necessary. The five stock structure hypotheses currently under consideration (**figs from SDWP1 or Bryde's whale IST docs**) concern the extent of mixing between different putative subpopulations in the different sub-areas. The range of hypotheses, and indeed the sub-area boundaries, are not set in stone, so there are maybe three distinct tasks to consider:

- to use genetic data to evaluate the relative plausibility of the various hypotheses currently under consideration;
- to use genetic data to help parametrize quantities such as mixing proportions (of breeding sub-populations in feeding sub-areas) and dispersal proportions (migration between sub-populations);
- to consider whether the range of hypotheses needs alteration/expansion.

Equations for harvesting under dispersal example in Annex I, and link to genetic differentiation parameters

The relationship between harvesting, dispersal rates and conservation goals can be explored using the following simple population dynamics model for two populations linked by dispersal.

$$N^1_{y+1} = N^1_y + rN^1_y(1-N^1_y/K^1) - C_y - \phi^{1\to 2}N^1_y + \phi^{2\to 1}N^2_y$$
$$N^2_{y+1} = N^2_y + rN^2_y(1-N^2_y/K^2) + \phi^{1\to 2}N^1_y - \phi^{2\to 1}N^2_y \qquad (1)$$

where N^i_y is the number of animals in stock *i* at the start of year *y*,

r is the intrinsic rate of growth (assumed to be 0.1),

K^i is the carrying capacity of stock *i*:

$$K^1 = \alpha K^{\text{tot}}; K^2 = (1-\alpha)K^{\text{tot}} \qquad (2)$$

K^{tot} is the total carrying capacity (assumed to be 1000),

α is the proportion which stock 1 is of the total at pre-exploitation equilibrium,

$\phi^{i\to j}$ is the proportion of animals of stock *i* which migrate (permanently) each year to stock *j*; by definition:

$$\phi^{2\to 1} = \phi^{1\to 2}K^1/K^2 \qquad (3)$$

C_y is the catch during year *y*, assumed to be taken only from stock 1, and computed by applying an exploitation rate λ to the total population size (both stocks combined) at the start of year *y*, i.e.:

$$C_y = \lambda(N^1_y + N^2_y) \qquad (4)$$

The populations do not have the same carrying capacities and only one of the populations is subject to harvest, but the catch limit for population 1 is (somewhat unintentionally) based on the total numbers in both populations (because it is not known that there two populations). Note that the per capita dispersal rates at equilibrium must be different for the two populations, in order to maintain different population sizes.

For the calculations in Annex I Figure 1, the exploitation rate λ is set to 0.05, corresponding to *MSY* in the panmictic case (i.e. very high dispersal). In that case, the equilibrium abundance relative to unexploited would be 50%.

Translation between the dispersal rate phi and genetic differentiation parameters might be accomplished as follows. If plausible values for the N/Ne ratio and the generation time are available, per capita annual dispersal can be easily transformed into genetic dispersal rates (m*Ne). Under mutation-drift equilibrium and a specified migration model, there is a simple formula relating m*Ne to Fst or related quantities (Wright, 1978). However, under non-equilibrium scenarios or other violations of the rather stringent underlying assumptions, this formula is biassed, and more sophisticated methods (e.g. coalescence) would be recommended for analysis of real data. To what extent this extra sophistication would be worthwhile for purposes of setting a “target Fst” will depend on the specific context being considered.