S1: Detailed Genetic Methodology

Microsatellite analysis

Thirteen polymorphic microsatellite loci [EV1, EV5, EV37, EV94, EV104 (Valsecchi and Amos 1996); SW10, SW19 (Richard et al. 1996); FCB1, FCB14, FCB17 (Buchanan et al. 1996); GATA28, GATA417 (Palsbøll et al. 1997); TEXVET5 (Rooney et al. 1999)] were amplified in separate 15 ul PCR reactions with the following conditions: 100 mM Tris HCl, pH 8.4, 500 mM KCl, 1.5 to 2.0 mM $MgCl_2$, 200 uM of each dNTP, 0.25uM of each primer, 0.025uM of a labelled primer, 0.3 units of Taq, and 10-100 ng of template DNA. Two thermocycling profiles (denoted as PM1 and PM2) were utilized for all loci tested. The PM1 profile consisted of an initial denaturing step of 95ºC for five minutes, 35 cycles (30 seconds (s) at 95ºC, 60 s at specified annealing temp, 60 s at 72ºC) followed by a final extension step of eight minutes at 72ºC. The following 13 loci and their respective annealing temperatures were run in the PM1 profile: EV1: 57 ºC, EV5: 59 ºC, EV94: 55 ºC, EV104: 54 ºC, SW10: 56 ºC, SW19: 56 ºC, FCB1: 53 ºC, FCB14: 53 ºC, FCB17: 56 ºC, GATA28: 53 ºC and GATA417: 56 ºC. The PM2 profile consisted of an initial denaturing step of 95ºC for five minutes, 35 cycles (60 seconds (s) at specified annealing temp, 60 s at 72ºC, 40 s at 94ºC) followed by one cycle at the specified annealing temp and a final extension step of five minutes at 72ºC. The following two loci and their respective annealing temperatures were run in the PM2 profile: EV37: 56 ºC and TEXVET5: 60 ºC. Amplified DNA was analysed for length variation on 6% polyacrylammide denaturing gels using fluorescent imaging on an automated ABI PRISM 377 (Applied Biosystems, Warrington, UK). Results were then imported into the program Genotyper 2.0^{TM} so that allele sizes could be visualized and

scored. Samples were defined as unique individuals by comparing alleles at six to ten of the least variable loci by eye (followed by confirmation using the CERVUS 2.0 program), in addition to incorporating mtDNA haplotype and gender data for each individual sample (when possible). Paetkau and Stroebek's (Paetkau and Strobeck 1994) probability of identity method was utilized to provide assurance that duplicate samples have been removed. The resulting genotypes were used to provide an analysis of kinship structure based on bi-parentally inherited nuclear DNA.

The number of unique alleles was calculated for each locus and over all loci and the observed heterozygosity (H_O) and expected heterozygosity (H_E) levels at each locus were tested for deviation from HWE using Fisher's exact test in the GENEPOP 3.2a program (Guo and Thompson 1992). GENEPOP 3.2a tests the probability of Fisher's exact test using the Markov chain method with 1000 dememorizations, 1000 batches and 1000 iterations. The frequency of putative null alleles for all 13 loci/population was tested using the computer program CERVUS 2.0 (Marshall et al. 1998). A test for linkage disequilibrium (null hypothesis: independence between genotypes at separate loci) was completed for each pair of loci using GENEPOP 3.2a to determine whether associations existed between pairs of alleles using Fisher's exact test and the Markov chain method (1,000 dememorizations, 1000 batches and 1000 iterations per batch were used for each population).

mtDNA analysis

The first 399 bp at the 5'-end of the mtDNA control region was amplified for 219 samples using primers L15812 5'-CCTCCCTAAGACTCAAGG-3' (Arnason et al. 1993) and H16343 5'-CCTGAGAATGCAACTAGAGG-3' (Southern et al. 1988) in separate

30 ul PCR reactions with the following conditions: 100 mM Tris HCl, pH 8.4, 500 mM KCl, 1.3 mM $MgCl_s$, 200 uM of each dNTP, 0.25 uM of each primer, 0.3 units of Taq, and 10-100 ng of template DNA. The PCR thermocycling profile consisted of an initial denaturing step of 95ºC for five minutes, 35 cycles (90 seconds (s) at 55ºC, 90 s at 72ºC, 45 s at 72ºC), followed by one cycle of 55º for 90 s, and a final extension step of eight min at 72ºC. Purified PCR product was sequenced directly using the ABI 377 and BigDye sequencing kit (Applied Biosystems). Sequences were automatically aligned and then edited by eye using Sequencher 4.2.1 software (Gene Codes Corp. Ann Arbor, MI, USA). Both strands (forward and reverse) were sequenced for unique haplotypes as a means of polymorphic site confirmation. Sequence alignment for multiple individuals was done automatically using Sequencher 4.1.2 software. Multiple sequence alignments and their corresponding chromatograms were edited by eye for discrepancies in base-calling. The resulting sequences were aligned with twenty-three sperm whale sequences provided by Dr. Sarah Mesnick (Southwest Fisheries Science Centre) using the program Mega 2.0 (T. Kumar, K. Tamura, M. Nei, Pennsylvania State University, http://www.megasoftware.net/) , several of which are described in Lyrholm and Gyllensten (Lyrholm and Gyllensten 1998).

Gender Determination

Gender determination was performed using odontocete-specific primers (ZFYX0582F, ZFY0767F, and ZFX0923R) that amplify the ZFX and the ZFY sequence (Bérubé and Palsbøll 1996). All sperm whale samples were tested (including strandings) to determine and confirm gender using the techniques described above. Amplification of the ZFX and

ZFY fragments were carried out in 20 PCR µl reactions with the following conditions: 67 mM Tris-HCl, pH 8.8, 2.0 mM $MgCl_2$, 16.6 mM $(NH_4)_2SO_4$, 10 mM β-mercaptoethanol, 200 µM of each dNTP, 0.5 µM of each primer, 0.4 units of Taq™ DNA polymerase and 100 ng of template DNA. Amplification reactions were run on a MJ Research, INC. PTC-100™ programmable thermal controller with an initial 5-min denaturing step at 94ºC, 37 cycles (60 s at 94ºC, 60 s at 52ºC, 60 s at 72ºC) followed by a 5-min final extension at 72ºC. Amplification products were distinguished on a 1.2% Agarose gel and scored against a DNA 100 bp ladder (Sigma). Females were defined as having only one band at approximately 383 bp, while males possessed a characteristic 227 base pairs (bp) fragment in addition to the 383 bp fragment. It is this 383 bp fragment that acts as a positive control verifying that the amplification reaction has taken place. Four controls, consisting of two confirmed male and two confirmed female samples taken from stranding events in the North Sea and Gulf of Mexico respectively, were run with each set of unknown samples to insure that both proper amplification had occurred as well as to provide a means of gender verification. A negative control, composed of the stock solutions only, was run to check for potential cross-contamination. Each individual sample was molecularly sexed using the ZFY/ZFX method a minimum of two times to verify the results.

Relatedness and Kinship Assessments

Levels of genetic relatedness were calculated for pairs of individuals within clusters and groups using the programs Relatedness 5.0.2 and Kinship 1.3.1 according to Queller and Goodnight's (1989) method. Relatedness 5.0.2 and Kinship 1.3.1 were used to estimates

Grafen's relatedness coefficient (*R*) between all pairs of individuals based on the number of shared microsatellite alleles and population allele frequencies using a regression measure of relatedness:

$$R = \frac{\sum\sum\sum x\,k\,l(P_y - P^*)}{\sum\sum\sum x\,k\,l(P_x - P^*)}$$

where *x* indexes individuals in the data set, *k* indexes loci and *l* indexes allelic position (i.e. *l* = 1 or 2 for a diploid individual, 1 only for a haploid). P_x and P_y are the frequency of the alleles within the current *x* and *y* individuals respectively. P^* is the frequency of the allele in the overall population (excluding all putative relatives of *x*) – determined from methods previously described. R-estimates are normally based on background population allele frequencies. As a result, R-values may be biased as common alleles are more likely to be shared by chance than descent. Kinship 1.3.1 is able to compensate for this type of bias by assigning lower R-values to individuals that share common alleles and higher R-values to those that share rare alleles. In the case of the GOS, an adequate sample size to build background allele frequencies on is not available from the Caribbean Sea population. Given the lack of nuclear DNA variation between populations across oceans (Lyrholm et al. 1999), the neighboring Gulf of Mexico population, which has been thoroughly analysed with respect to nuclear DNA and mtDNA, was used to provide the background allele frequencies needed to base relatedness estimates for the GOS. Relatedness values were based on background allele frequencies with highly related ($R \geq 0.40$) whales previously removed from the dataset (see (Engelhaupt 2004)). Relatedness measurements ranged from -1.0 to +1.0 with positive values signifying two individuals sharing more alleles that were identical by descent than expected by chance, whereas

negative R-values were indicative of two individuals sharing fewer such alleles than expected by chance. When populations are in Hardy-Weinberg equilibrium, relatedness coefficients should average 0.50 for first-order relatives (e.g. parent-offspring and full-sibling pairs), 0.25 for second-order relatives (e.g. half-sibs, grandparent-grandchild, aunt/uncle-niece/nephew) and 0.00 for pairs of randomly chosen individuals that are not related.

Kinship tests hypotheses of pedigree relationships between pairs of individuals by calculating a likelihood ratio that a pair of genotypes fits a particular hypothesized relationship (Goodnight and Queller 1999). The likelihood ratio is based on R-values, population allele frequencies and the pair's genotypes. For example, if we want to examine whether a pair of individuals with an R-value of 0.50 is a first-order relation, we set our null hypotheses (no relation) to zero and our primary hypothesis (first-order relation) to R=0.50. Log likelihood values are calculated as Kinship performs a simulation routine (set at 10,000 repeats) to determine a distribution of log likelihoods and significance levels for likelihood ratios using the hypothesis settings and the population allele frequencies. Rejection or non-rejection of the null and primary hypotheses was determined based on the resulting significance levels.

References:

Arnason U, Gullberg A, Widegren B (1993) Cetacean mitochondrial DNA control region: sequences for all extant whales and two sperm whale species. Mol. Biol. Evol. 10:960-970

Bérubé M, Palsbøll P (1996) Identification of sex in cetaceans by multiplexing with three ZFX and ZFY specific primers. Mol. Ecol. 5:283-287

Buchanan FC, Friesen MK, Littlejohn RP, Clayton JW (1996) Microsatellites from the beluga whale *Delphinapterus leucas*. Mol. Ecol. 5:571-575

Engelhaupt D (2004) Phylogeography, kinship and molecular ecology of sperm whales (*Physeter macrocephalus*). Ph.D. Thesis, Department of Biological Sciences, University of Durham. 191pp

Goodnight KF, Queller DC (1999) Computer software for performing likelihood tests of pedigree relationships using genetic markers. Mol. Ecol. 8:1231-1234

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportions for multiple alleles. Biometrics 48:361-372

Lyrholm T, Gyllensten U (1998) Global matrilineal population structure in sperm whales as indicated by mitochondrial DNA sequences. Proc. R. Soc. Lond. B. 265:1679-1684

Lyrholm T, Leimar O, Johanneson B, and Gyllensten U (1999) Sex-biased dispersal in sperm whales: contrasting mitochondrial and nuclear genetic structure of global populations. Proc. R. Soc. Lond. B. 266:347-354

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7:639-655

Paetkau D, Strobeck C (1994) Microsatellite analysis of genetic variation in black bear populations. Mol. Ecol. 3:489-495

Palsbøll PJ, Bérubé M, Larsen F, Jorgensen H (1997) Primers for the amplification of tri- and tertamer microsatellite loci in baleen whales. Mol. Ecol. 6:893-895

Richard KR, Whitehead H, Wright JM (1996) Polymorphic microsatellites from sperm whales and their use in the genetic identification of individuals from naturally sloughed pieces of skin. Mol. Ecol. 5:313-315

Rooney AP, Merrit DB, Derr JN (1999) Microsatellite diversity in captive bottlenose dolphins (*Tursiops truncatus*). J. Hered. 90:228-231

Southern SO, Southern PJ, Dizon A (1988) Molecular characterization of a cloned dolphin mitochondrial genome. J. Mol. Evol. 28:32-42

Valsecchi E, Amos W (1996) Microsatellite markers for the study of cetacean populations. Mol. Ecol. 5:151-156