

Annex N

Report of the Working Group on DNA

MEMBERS: Pastene (Chair), Andersen, Cipriano, Corkeron, Double, Goto, Gregovich, Hoelzel, Jorde, Kakuda, LeDuc, Morin, Park, Skaug, Strand, Tiedemann, Vikingsson, Waples.

1. ELECTION OF CHAIR

Pastene convened and chaired the group.

2. APPOINTMENT OF RAPORTEURS

LeDuc acted as rapporteur.

3. ADOPTION OF AGENDA

The agenda is given as Appendix 1. Items 5, 6 and 7 of the Agenda are in response to requirements placed on the Scientific Committee by IWC Resolution 1999-8 (IWC, 2000), which called for annual reports on progress in the following areas:

- (1) Genetic methods for species, stocks and individual identification.
- (2) Collection and archiving of tissue samples from catches and bycatch.
- (3) Status of and conditions for access to reference databases of DNA sequences or microsatellite profiles derived from directed catches, bycatch, frozen stockpiles and products impounded or seized because of suspected infractions.

4. REVIEW OF DOCUMENTS

Relevant information was contained in SC/59/SD1, SC/59/SD2 and SC/59/SD5.

5. PROGRESS ON GENETIC METHODS FOR SPECIES, STOCK AND INDIVIDUAL IDENTIFICATION

SC/59/SD1 describes the testing of DNA extraction and PCR amplification methods developed for human forensic analysis for use with degraded, damaged, and highly processed cetacean tissues, and including: (1) a series of purification steps that can purify degraded and chemically treated DNA from “processed” samples, (2) pre-amplification or “primerless PCR” which repairs and improves DNA template material prior to PCR amplification, and (3) hemi-nested PCR which uses hierarchical PCR to amplify cetacean DNA with high specificity, even in a background containing degradation products and other inhibitors and DNA from other species. Initial trials using some of these new approaches were much better able to amplify, and then precisely identify, DNA sequences from highly degraded materials of several types, including decomposed and mummified dolphin carcasses and partially digested blubber recovered from a sleeper shark stomach. Although some products could be analyzed with the addition of only one of these improvements, the most challenging sample (blubber from the shark stomach) required application of the full 3-step extraction/purification protocol, pre-amplification and hemi-nested PCR to produce a usable product.

In discussion Morin noted that the paper advocated the use of phenol/chloroform DNA extraction methods, but that some ancient DNA labs now use silica-based methods, which yield smaller but cleaner amounts of DNA. He informed the Group about a recent published paper on ancient DNA extraction (Rohland and Hofreiter, 2007). Hoelzel asked what proportion of the samples were ancient and what proportion were recent but difficult. He noted that in some cases, the important step is getting rid of inhibitors, in which case DNA purification columns can work well. Other options include stepwise purification and trying different types of Taq polymerase. In response Cipriano noted that all samples tested were difficult but not ancient, and pointed out that the use of proof-reading and high-fidelity Taq polymerases was also helpful in working with degraded or highly-processed specimens and also noted that bone demineralization techniques now being adopted by the human forensics

community will also be tested in further controlled experiments with cetacean bone and tooth samples. In response to a question Cipriano responded that the typical length of sequences obtained is 120bp.

SC/59/SD2 presented preliminary estimates of genotyping error rates in the Norwegian minke whale DNA-register using DNA-profiles from 589 mother-foetus pairs. The basic idea is that mother and offspring must share at least one allele per locus. It was reported that the laboratory currently used for the DNA-register has a much lower error rate than the laboratory used until 2002. This conclusion is supported by auxiliary data consisting of a repeated scoring of 25 individuals for which the true genotype is believed to be known. The error rates for the period 2002-present are comparable to those found in the published literature. The error rate per allele ranged between 0.0016 and 0.18, depending on the dataset and the assumptions made.

It was noted that the analytical aspects of this document were discussed in the Working Group on Stock Definition. Only a brief discussion of this document took place here. The issue of differences in 1bp error was raised. Tiedemann noted that the type of microsatellite used would be useful information to evaluate this. A 1bp error is rare for 4bp microsatellite repeats, unless it reflects an even rarer mutation in the flanking region. Hoelzel pointed out that these errors can also result from adenylation. Waples noted that the focus on scoring errors ignores actual mutations. Skaug responded that there have been cases where mutations have been indicated.

SC/59/SD5 described the improved and expanded *Witness for the Whales, Vs 4.3* database and accompanying *DNA Surveillance* web-based program for species identification using DNA sequence data and phylogenetic analysis (available at www.dna-surveillance.auckland.ac.nz). The *Witness for the Whales Vs3.1* database has been revised by replacing all sequences from specimens of unknown provenance with sequences from known-provenance individuals. *Witness for the Whales, database Vs 4.3* is now taxonomically comprehensive, with a total of 399 control region sequences and 264 cyt *b* sequences representing 88 species. Sequences from documented specimens now represent all of the 83 species recognized by Rice (1998), with two exceptions: the Atlantic hump-backed dolphin, *Sousa teuszii* and the Indian hump-backed dolphin *S. plumbea* (the latter of which has not been accepted by IWC). Vs 4.3 also includes seven species proposed in recent publications and three subspecies of baleen whales. A total of 47 new control region sequences have been submitted to GenBank, and all sequences in the cytochrome *b* dataset are already available there. The aligned sequences for each of the eight taxonomically organized subsets of sequences can be downloaded from *DNA Surveillance* for further user-based analyses. The database now includes sequences from seven new species proposed in recent publications and three subspecies of baleen whales. Both the control region and cytochrome *b* datasets include unique sequences from 2-6 specimens for most species. Software improvements to *DNA Surveillance* in Vs 4.3 include improved security, delegated administration of reference datasets for improved curatorial access, administrator control of download permissions of aligned test and reference sequences, and preliminary reference databases for six other taxonomic groups including sea horses, a group recently placed under control of the Convention on International Trade in Endangered Species (CITES).

Last year the Committee agreed on several tasks to be conducted intersessionally (IWC, 2007, p. 57) to continue with the development of plans for sequence validation, specifically:

1. Determine how many sequences exist in GeneBank for the baleen whale species currently under genetic investigation by the Committee;
2. How many are typically added annually;
3. The expertise and time required to carry out the first round of validation and subsequent annual rounds, and
4. The approximate cost if done under contract.

Appendix 2 summarizes the intersessional work and proposes a mechanism to start the first round of validation. A search in GenBank revealed that by April 2007 a total of 1,323 mtDNA control region sequences of baleen whales were deposited in this sequence depository. There has been considerable variation in the number of sequences deposited per year. Appendix 1 proposes to start with the first round of validation using *DNA Surveillance* (Ross *et al.*, 2003; Ross and Murugan, 2006) in the intersessional period 2007/08 for the sequences deposited prior to 2007 (n= 922), under contract with Dr. Ross (Auckland University, New Zealand). The Working Group strongly **agreed** with the proposal. The Group also **agreed** that the validation would take the form of a report with the provisions specified in Appendix 1.

As agreed by the Committee last year, any anomaly detected in the validation process would be shared with members of the Committee. The original submitter would be notified of the inconsistency and a suggestion would be made that an amendment be made to the entry, followed with an offer to help after *GenBank* has been notified directly of the permission to amend. The Group **noted** that a member of the Committee should be

identified to carry out this work after the report of the first round of validation is received (2008 Committee meeting).

The possibility of future annual validation (after 2008) to be made under contract will be discussed after the experience of the first round of validation has been evaluated by the Working Group.

6. PROGRESS ON COLLECTION AND ARCHIVING OF SAMPLES FROM CATCHES AND BYCATCHES

Skaug reported on the status of the Norwegian register (Appendix 3). The collection of samples is from the commercial catches of common minke whales from 1997 to 2006. The number of samples missing from the register by year ranged from 0-11. Some of the missing samples reflect unsampled whales, while others resulted from inadvertent duplicates. It was noted that there was an unusually high number of duplicate samples (7) from 2004, which coincided with the end of the use of government inspectors for handling samples.

Goto reported on the status of the Japanese register (Appendix 4). The collection of samples is from scientific whaling in the Antarctic (JARPA-JARPA II) and North Pacific (JARPN II), bycatches and strandings. It includes complete coverage for 2006 through the 2006/2007 Antarctic season.

Vikingsson reported on the status of the Icelandic register (Appendix 5). Samples are presently in hand for all whales taken in 2003-2006.

7. REFERENCE DATABASES AND STANDARDS FOR A DIAGNOSTIC REGISTER OF DNA PROFILES

Genetic analyses have been completed and data on mtDNA, STRs and sex entered in the Norwegian register for years through 2004 (Appendix 3).

For the Japanese register (Appendix 4), all analyses have been completed for NP minke, NP Bryde's and NP sei whales through 2006, mtDNA for NP sperm whales through 2006 and sex for all samples from all species. The genetic samples of Antarctic minke whales and southern fin whales have not been analyzed yet, except for sex. For bycatch samples, mtDNA and sex have been completed for all samples through 2006, and mtDNA for the stranding samples for 2006.

Vikingsson noted that the Icelandic register (Appendix 5) contains the same items in the same format as the Norwegian register so that cross-comparisons can be carried out. All genetic analyses have been conducted for fin whales caught in 2006.

Regarding the report on the status of the registers there was also discussion about the inclusion of background information with the register updates for the benefit of new members of the Group. The Chair reminded the Group that specifications of the registers in Norway and Japan had been documented previously. He suggested that report of updates of registers should include a list of references including these relevant documents. The Group **agreed** with the Chair.

It was also suggested that new technological improvements in the registers, including the development of new microsatellite markers, could be reported in the annual report of the status of the registers. The Chair suggested that any substantial new technical improvement in the registers should instead be presented to the Group as separate papers. The Group **concurred** with his suggestion.

8. WORK PLAN

The terms of reference for the Working Group for the next year will remain the same as for this year, unless the Commission requests other information in the interim. Members of the Working Group are encouraged to submit papers relating to these terms of reference and to propose additional agenda items. In particular the Group looks forward for the presentation and discussion of the recent published paper by Rohland and Hofreiter (2007). The Group **agreed** that the first round of sequence validation is important and **recommended** that it should be conducted in the intersessional period. The budgetary implication for this work is discussed under Item 21 of the Committee report.

9. ADOPTION OF REPORT

The report was adopted by consensus.

REFERENCES

- International Whaling Commission. 2000. Chairman's Report of the Fifty-First Annual Meeting. Appendix 9. IWC Resolution 1999-8. Resolution on DNA testing. *Rep. Int. Whaling Commn* 1999:55.
- International Whaling Commission. 2007. Report of the Scientific Committee. *J. Cetacean Res. Manage.* 9 (Suppl.):1-302.
- Rice D.W. 1998. Marine mammals of the world: systematics and distribution, Special Publication Number 4, The Society for Marine Mammalogy: Allen Press, USA. ix+231pp.
- Rohland, N. and Hofreiter, M.2007. Comparison and optimization of ancient DNA extraction. *BioTechniques* 42(3): 343-403.
- Ross, H. A., Lento, G. M., Dalebout, M. L., Goode, M., Ewing, G., McLaren, P., Rodrigo, A. G., Lavery, S., Baker, C. S. 2003. DNA Surveillance: Web-based molecular identification of whales, dolphins, and porpoises. *J. Hered.* 94:111-114.
- Ross, H. A. and Murugan, S. 2006. Using phylogenetic analyses and reference datasets to validate the species identity of cetacean sequences in GenBank. *Mol. Phyl. Evol.*40: 866-871.

Appendix 1

AGENDA

1. Election of Chair
2. Appointment of rapporteurs
3. Adoption of Agenda
4. Review of documents
5. Progress on genetic methods for species, stock and individual identification
6. Progress on collection and archiving of samples from catches and bycatches
7. Reference databases and standards for a diagnostic register of DNA profiles
8. Work plan
9. Adoption of report

Appendix 2

REPORT OF THE INTERSESSIONAL WORK ON VALIDATION ON DNA SEQUENCES IN GENBANK

Luis A. Pastene

Last year the Committee agreed to re-establish the intersessional Working group under Pastene to continue with the development of plans for sequence validation, specifically:

1. Determine how many sequences exist in GeneBank for the baleen whale species currently under genetic investigation by the Committee;
2. How many are typically added annually;
3. The expertise and time required to carry out the first round of validation and subsequent annual rounds, and
4. The approximate cost if done under contract.

REPORT OF THE INTERSESSIONAL WORK

Number of sequences in GenBank

Pastene contacted Dr. Howard Ross (Senior Lecturer, Bioinformatics Institute, School of Biological Sciences, University of Auckland, New Zealand). Dr. Ross is the author of *DNA Surveillance* (Ross *et al.*, 2003; Ross and Murugan, 2006), the available vehicle agreed by the Committee last year to carry out the sequence validation.

Dr. Ross provided information on the number of mtDNA control region sequences for baleen whales deposited at the *GenBank*, by year:

Pre-2000: 189
2000: 0
2001: 159
2002: 11
2003: 59
2004: 300
2005: 143
2006: 61
2007 (by April): 401

There has been considerable variation from year to year.

SEQUENCE VALIDATION AND APPROXIMATE COST

Dr. Ross offered his service to carry out the validation work under contract with the IWC. The number of sequences of baleen whales until 2006 is 922. The validation cost for such number would be 7,240 NZ dollars (2,682 pounds).

PROPOSAL FOR THE ROUNDS OF VALIDATION

Given the expertise involved the proposal is to accept the offer by Dr. Ross followed by annual contracts for validation on an annual (or biennial) basis (after the experience and results from the first round of validation has been evaluated by the Working Group). For example the results of the validation of the sequences until 2006 would be presented to the Committee during the 2008 meeting. During that meeting a new contract would be made for validation of the sequences deposited in 2007 with a corresponding report to be presented during the 2009 Committee meeting, and so on.

REPORTS OF THE VALIDATION ROUNDS

The Working Group would provide Dr. Ross with data and instructions which would enable him to successfully carry out the service specified. It is envisaged that the validation would take the form of a report with the following provisions:

1. List the *GenBank* accession number and species identity of each mysticete control region sequence with the species identity as determined using the most recent version of the Witness for the Whale reference sequence alignments (see SC/59/SD5) and the DNA Surveillance software engine.

2. The above list to be supported by phylogenetic trees, one per sequence, showing the placement of the *GenBank* sequence in relation to the reference sequence.
3. An evaluation of the types of inconsistencies/errors as agreed by the Committee last year: quality of submitted sequences, accuracy of species identification and accuracy of geographical location.

As agreed by the Committee last year, any anomaly detected in the validation would be shared with members of the Committee. The original submitter would be notified of the inconsistency and a suggestion would be made that an amendment be made to the entry, offering to help after *GenBank* has been notified directly of the permission to amend. A member of the Committee should be identified to carry out this work.

REFERENCES

- Ross, H. A. and Murugan, S. 2006. Using phylogenetic analyses and reference datasets to validate the species identity of cetacean sequences in GenBank. *Mol. Phyl. Evol.* 40: 866-871.
- Ross, H. A., Lento, G. M., Dalebout, M. L., Goode, M., Ewing, G., McLaren, P., Rodrigo, A. G., Lavery, S., Baker, C. S. 2003. DNA Surveillance: Web-based molecular identification of whales, dolphins, and porpoises. *J. Hered.* 94:111-114.

Appendix 3

STATUS OF THE NORWEGIAN MINKE WHALE DNA-REGISTER BY FEBRUARY 2007

Hans Julius Skaug

year	DNA-register ¹	IWC catch statistics ²	not landed ³	landed ⁴	duplicates ⁵	missing samples ⁶	total missing ⁷
1997	488	503	7	496	3	5	8
1998	609	625	11	614	1	4	5
1999	571	591	17	574	2	1	3
2000	470	487	6	481	3	8	11
2001	538	552	11	541	2	1	3
2002	625	634	9	625	0	0	0
2003	637	647	9	638	1	0	1
2004	530	544	7	537	7	0	7
2005 ⁸	-	639	6	633	-	4	-
2006 ⁹	-	545	7	538	-	2	-

The number of individuals contained in the DNA-register, and the number of individuals missing. For 2005 and 2006 the genetic analyses are not completed, as indicated by the ‘-’ in the table.

¹ Number of unique individuals contained in the DNA-register (not containing duplicates).

² Number of individuals caught by Norway, including individuals not landed.

³ Number of individuals killed, but not taken onboard the vessel.

⁴ Number of individuals taken onboard the vessel.

⁵ Number of occurrences of (tissue) sample switching on board the vessel as detected by comparison of genetic profiles. The result is that two samples have been returned from one individual, and no sample has been returned for one individual.

⁶ Number of individuals for which tissue samples are missing for other reasons than sample switching.

⁷ The difference between the columns “Landed” and “DNA-register”.

⁸ Genetic analyses not yet completed.

⁹ Tissue samples collected, but not yet sent to genetic laboratory.

¹⁰ Olaisen, B. (1997) Proposed specifications for a Norwegian DNA database register for minke whales. Paper SC/49/NA1 presented to the Scientific Committee of the IWC.

Appendix 4

AN UPDATE OF THE JAPANESE DNA REGISTER FOR LARGE WHALES

Mutsuo Goto and Naohisa Kanda

Source/Species	Period	Genetic	mtDNA	STRs	Sex	
			samples			
Scientific whaling						
NP minke whale	05	220	220	100	220	
		06	195	195	100	195
NP Bryde's whale		05	50	50	50	50
		06	50	50	50	50
NP sei whale	05	100	100	100	100	
		06	100	100	100	100
NP sperm whale	05	5	5	0	5	
		06	6	6	0	6
Antarctic minke whale	05/06	853	0	0	853	
		06/07	505	0	0	505
Antarctic fin whale	05/06	10	0	0	10	
		06/07	3	0	0	3
By-catches						
NP minke whale	06	147	147	0	147	
NP humpback whale	06	3	3	0	3	
NP sperm whale	06	1	1	0	1	
Strandings						
NP minke whale	06	8	8	0	0	
NP Bryde's whale		06	3	3	0	0
NP sperm whale	06	2	2	0	0	
NP humpback whale	06	1	1	0	0	
NP right whale	06	1	1	0	0	

STR = microsatellites, NP = North Pacific.

Note 1: sex of the whales taken by scientific whaling was determined by scientists onboard the research vessels.

Note 2: 0 = not yet analysed at the time this report was prepared.

*Pastene, L.A. and Goto, M. Status of the Japanese DNA register for large whales. *J. Cetacean Res. Manage.* 8 (Suppl.): 255-258.

Appendix 5
STATUS OF THE ICELANDIC MINKE WHALE DNA REGISTER

Gisli A. Vikingsson

Practical arrangements regarding the establishment of the Icelandic DNA register were concluded earlier this year (2007). The Marine Research Institute, Reykjavik, will be responsible for the establishment and maintenance of the registry that will be of the same format as the Norwegian DNA registry.

Table 1 gives the present status of the registry. Samples from all the common minke whales landed as a part of the Icelandic research program to date (2003-2006) as well as from commercial catches of one minke whale and seven fin whales have been archived. Genetic analyses have been completed for fin whales and those for common minke whales are at a final stage.

Table 1

Species year	Type ¹	No. genetic samples	Microsatellites	MtDNA	Sex
Common minke whales					
2003	SP	36	0	0	0
2004	SP	25	0	0	0
2005	SP	34	0	0	0
2006	SP	58	0	0	0
2006	C	1	0	0	0
Fin whales					
2006	C	7	7	7	7

¹SP= Special Permit Catch; C= Commercial Catch.