Report of the Scientific Committee

Bled, Slovenia, 7-19 June 2016

Annex O: Report of the Working Group on DNA

**This report is presented as it was at SC/66b.
There may be further editorial changes (e.g. updated references, tables, figures) made before publication.**

International Whaling Commission
Bled, Slovenia, 2016

Annex O

Report of the Working Group on DNA

Members: Pastene (Chair), Baird, Baker, Bickham, Cipriano, DeWoody, Hrabkovsky, Hoelzel, Kontzen, Lang, Lee, Øien, Pampoulie, Park, Skaug, Tiedemann, Torres-Flores, Tsuji, Waples, Yoshida

1. ELECTION OF CHAIR
Pastene convened and chaired the Group.

2. APPOINTMENT OF RAPPORTEURS
Pastene acted as rapporteur assisted by Tiedemann.

3. ADOPTION OF AGENDA
The adopted agenda is given as Appendix 1. Items 5, 6, 7 and 8 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
Documents SC/66b/DNA01-04 were relevant for the Working Group.

5. PROGRESS ON GENETIC METHODS FOR SPECIES, STOCK AND INDIVIDUAL IDENTIFICATION
SC/66b/DNA01 responded to the recommendation from the JARPNII final review workshop that genotyping error rates should be estimated. For this aim, a total of 200 common minke whales from JARPN/JARPNII (approximately 8% of the total available samples) were randomly selected and newly genotyped at the same16 microsatellite loci (repeat-genotyping). The genotyping error rate combined over all loci and all samples was low, 0.0044 per reaction or 0.0025 per allele. These rates were similar to the rates estimated for fur seals and lower than the rates estimated for bowhead whales.

The Working Group noted that this type of error estimates does not measure accuracy (i.e., the correct typing of a true allele size), but precision (aka consistency, i.e., repeatability of genotyping). It was further noted that this is a general issue regarding typing error estimates (see IWC SC data quality guidelines for details). It was confirmed that this paper measures the genotyping error in the sense recommended by the data quality guidelines. Therefore the Working Group **agreed** that the work presented in SC/66b/DNA01 addresses this recommendation made by the JARPNII review workshop appropriately.

SC/66b/DNA02 informed the Norwegian plan to update its DNA register. The Norwegian Minke Whale DNA Register (NMDR) is at present based on genotyping microsatellites, a single sex marker and sequencing the mtDNA control region. SC/66b/DNA02 presented plans to upgrade the NMDR by genotyping a suite of carefully selected SNPs which will still keep the register's primary function of traceability of whale products in Norway and the international market.

The Working Group welcomed Norway's plan to add SNPs in its register and noted that SNPs genotyping should be seen as a complement, not as a replacement of the current microsatellites genotyping. SNPs should be identified carefully. No technical details of the plan were available in SC/66b/DNA02 therefore the Working Group was unable to provide technical advices.

SC/66b/DNA03 reports a pilot study of a double digest RAD (ddRAD) protocol in Blainville's and Cuvier's beaked whales. Four samples from each species were run, with all samples from Blainville's coming from El Hierro, Canary Islands, and the Cuvier's samples coming from the Canary Islands, Scotland and the Mediterranean. The pilot study produced 9.2M quality controlled reads for the Blainville's and 16.4M quality controlled reads for the Cuvier's beaked whales. After loci construction and filtering in program STACKS, this produced 8,143 variable RAD loci for Blainville's and 14,095 variable RAD loci for Cuvier's beaked whales at moderate depths (20x). The higher variability in Cuvier´s beaked whales was probably due to the difference in sequencing success between the species and the broader geographic range of the Cuvier's compared with the Blainville's samples. The data were also analysed using PYRAD to identify loci in common across the two species; this revealed 9666 loci at 20x depth in common between at least one sample per species.

The study in SC/66b/DNA03 was considered a valuable proof-of-principle by the Working Group. The Group noted, however, that loci were compared across different genera. Therefore, the loci shared across the analysed species may not necessarily be considered orthologous (i.e., homologous and positioned at the same site in the genome).

SC/66b/DNA04 provided the first description of the gray whale genome and characterised a novel SNP panel that includes 88 gene-associated markers, two molecular sexing markers, and two mitochondrial markers. One male and one female western gray whale, and one female eastern gray whale were sequenced. Approximately 22,000 genes, a number similar to other cetacean genomes, were annotated. The gray whale is only the third species of baleen whales to have a genome sequence. Molecular markers such as single nucleotide polymorphisms (SNPs) can reveal otherwise cryptic aspects of organismal ecology and evolution. SC/66b/DNA04 sequenced the gray whale genome, repeatedly genotyped replicate whale biopsies at 92 SNP loci, then quantified genotyping error rates and variability at each marker. Mitochondrial DNA haplotyping and molecular sexing with SNPs was 100% concordant with conventional assays based on PCR and dideoxy sequencing or electrophoresis. Genotyping error rates, calculated across loci and across replicate samples, were very low (0.021%) and observed heterozygosity was 0.33 averaged over all autosomal markers. This level of variability across loci provides substantial discriminatory power, as evidenced by the genetic documentation of parent/offspring pairs in the study. For example, the mean probability of identity was $<10^{-25}$ for unrelated individuals and the mean probability of exclusion was >0.9999 when neither parent was known. The characterisation of the gray whale genome should enable comparative studies of natural selection in cetaceans and the SNP markers should be highly informative for future studies of gray whale population structure, demography and relatedness.

The output of the study was considered valuable for forensic applications in the context of the Working Group work. It was noted that – if SNPs occur in genes under strong selection – this selection need not necessarily be due to the SNP. If there is (positive) selection on a SNP, such SNP position is interesting to study divergence, and may serve well as a marker for forensic applications. It is however not applicable for any quantitative measure assuming selective neutrality (as many population parameters do). There was also some discussion on random sampling in a small population, like western North Pacific Gray whales. It was noted that any non-random sampling with regard to close kin (in particular mother/fetus pairs) should be avoided. However, other (random) sampling of close kin simply because of small population size is both unavoidable and acceptable.

It was further noted that availability of genome information is very helpful for SNP development. An alternative to the approach of comparing two full genomes (as used in this study) would be SNP identification by mapping of ddRAD sequences on a single genome.

6. REVIEW RESULTS OF THE 'AMENDMENTS' OF SEQUENCES DEPOSITED IN GENBANK

Last year the Committee encouraged Cipriano to keep contact with NCBI (National Center for Biotechnology Information) in the intersessional period to make progress on the mechanism for taxonomy updates at the NCBI In particular on the mechanism identified last year to allow annotation of *GenBank* sequences by interested parties, in order to note taxonomic mis-assignment or questions about geographic source of the organism involved (IWC, 2016 p71).

Cipriano informed the Working Group that although he did not correspond with Scott Federhen at NCBI in the past intersessional period, there was a new publication (Federhen, 2015) that acknowledged that there are misidentified sequences in *GenBank*, and entries with other annotation problems. It also noted that efforts should be spent to institute mechanisms to have these corrected (or flagged as problematic) to support the reference database requirement. These include inclusion of 'Sequence from type' which can help to alleviate these problems by providing a backbone of reliably identified sequence data. The Working Group **recommended** that cetacean holotype and paratype sequences should be archived in *GenBank* whenever possible (as has been done for the holotypes of *B. omurai* GenBank Accession No.AB201256 and *Mesoplodon perrini* Accession No AF441261).

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

The Committee previously endorsed a new standard format for the updates of national DNA registers to assist with the review of such updates (IWC, 2012 p53), and the new format worked well the last years. This year the update of the DNA registers by Japan, Norway and Iceland were based again on this new format.

Yoshida reported on the status of the Japanese register (Appendix 2). The collection of samples is from scientific whaling in the North Pacific (JARPN-JARPNII) and the Antarctic (JARPA-JARPAII), and from bycatches. It includes coverage for 1994-2015 (JARPN-JARPNII), 1987/88-2013/14 (JARPA-JARPAII). In the case of bycatches it includes coverage for 2001-2015.

Skaug reported on the status of the Norwegian register (Appendix 3). The collection of samples of North Atlantic common minke whale is from commercial catches for the period 1997 to 2015.

Pampoulie reported on the status of the Icelandic register (Appendix 4). The collection of samples is from scientific whaling and from commercial catches. It includes coverage for 2003-2007 (scientific whaling) and 2006-2015 (commercial catches).

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

An update of the Japanese register is shown in Appendix 2. For North Pacific minke whales sampled under JARPN II in 2015 mtDNA and microsatellite analyses of 100% (n=70) has been completed. For animals bycaught in 2015, mtDNA and microsatellite analyses has been completed for 100% (n=156).

For North Pacific Bryde's whales sampled under JARPNII in 2015, mtDNA and microsatellite analyses have been completed for 100% of the samples (n=25). No bycatch occurred in 2015. For North Pacific sei whales sampled under JARPNII in 2015, mtDNA and microsatellite analyses have been completed for 100% of the samples (n=90). No bycatch occurred in 2015. No sampling and bycatch of sperm whale occurred in 2015. No Antarctic minke whales were sampled in 2014/15.

For North Pacific humpback whales 100% of the ten whales bycaught in 2015 were screened for mtDNA and microsatellites. There was no bycatch of North Pacific right whales in 2015. No North Pacific fin whales were bycaught in 2015.

An update of the Norwegian register is shown in Appendix 3. After discounting for duplicates, missing samples and laboratory problems, 100% of the North Atlantic common minke whales caught in 2015 were screened for mtDNA and microsatellites (n=558).

An update of the Icelandic registry is shown in Appendix 4. For North Atlantic common minke whales 83% of the 29 animals taken in 2015 were screened for both mtDNA and microsatellites. 100% of the fin whales caught by commercial whaling in 2015 (n=154) were screened for both mtDNA and microsatellites.

The Group **appreciated** the efforts of Japan, Norway and Iceland in compiling and providing this detailed information of their registries.

9. OTHER

No other matters were discussed by the Group.

10. WORK PLAN

The terms of reference for the Working Group will remain the same for the next year, unless the Commission requests other information in the interim. Members of the Working Group were encouraged to submit papers relating to these terms of reference and to propose additional agenda items. Results of the 'amendment' work on sequences deposited in *GenBank* will be reported next year. Next year a comparison of methods for SNP development and assessment will be continued. Also the Working Group will examine the technical information relevant to the TORs of the Group, contained in documents presented to other groups and subcommittees.

11. ADOPTION OF THE REPORT

The report was adopted by consensus.

REFERENCES

Federhen, S. 2015. Type material in the NCBI taxonomy database. *Nucleic Acids Research* 43:

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. 2012. Report of the Scientific Committee. *J. Cetacean Res. Manage*. 13 (Suppl.): 1-87.

International Whaling Commission. 2016. Report of the Scientific Committee. *J. Cetacean Res. Manage*. 17 (Suppl.): 1-90.

Appendix 1

AGENDA

1. Election of Chair
2. Appointment of rapporteurs
3. Adoption of the Agenda
4. Review of documents
5. Progress on genetic methods for species, stock and individual identification
6. Review of results of the 'amendments' of sequences deposited in *GenBank*
7. Progress on collection and archiving of tissue samples from catches and bycatches
8. Reference databases and standards for a diagnostic DNA registries
9. Other
10. Working Plan
11. Adoption of the Report

Appendix 2

An update of the Japanese DNA register for large whales

Mutsuo Goto[1], Hiroyuki Oikawa[1] and Hideyoshi Yoshida[2]

[1] *The Institute of Cetacean Research, 4-5 Toyomi-cho, Chuo-ku, Tokyo, 104-0055, Japan*

[2] *National Research Institute of Far Seas Fisheries, 2-12-4 Fukuura, Kanazawa-ku, Yokohama 236-8648, Japan*

The status of the Japanese DNA register for large whales was presented and discussed during the 2005 IWC SC meeting (IWC, 2006). Since then, the number of genetic samples and the number of individuals analyzed and registered have been reported to the IWC SC annual meetings. The annual reports include information of whales taken by the scientific whaling in the North Pacific (JARPN/JARPNII) and the Antarctic (JARPA/JARPAII), and from bycatches and stranding. The most recent full description of the protocol used by the Institute of Cetacean Research for the genetic analyses in the context of the IWC guidelines was presented by Kanda *et al* . (2014).

The update of the Japanese DNA register for large whales till 2015 is as follows.

footnote #	1	2	3	4	5	6	7	8	9	10	11	12
Species/Year	type	# whales	# duplicate	# missing	# lab problem	#mtDNA	%mtDNA	#msat	%msat	sex analyzed	% sexed	note
NP minke whale												
1994-2014	SP	2573	0	0	8	2565	100	2565	100	2573	100	
2015	SP	70	0	0	0	70	100	70	100	70	100	
2001-2014	BC	1683	0	26	2	1683	100	1655	98	1653	98	
2015	BC	156	0	0	0	156	100	156	100	156	100	
NP Bryde's whale												
2000-2014	SP	680	0	0	3	677	100	680	100	680	100	
2015	SP	25	0	0	0	25	100	25	100	25	100	
2001-2014	BC	5	0	0	0	5	100	4	80	4	80	Include three Omura's whale
2015	BC	0	0	0	0	0	0	0	0	0	0	No BC.
NP sei whale												
2002-2014	SP	1174	0	0	4	1170	100	1174	100	1174	100	
2015	SP	90	0	0	0	90	100	90	100	90	100	
NP sperm whale												
2000-2014	SP	56	0	0	0	56	100	56	100	56	100	
2015	SP	0	0	0	0	0	0	0	0	0	0	No catch.
2001-2014	BC	2	0	0	0	2	100	2	100	2	100	
2015	BC	0	0	0	0	0	0	0	0	0	0	No BC.
Ant. minke whale												
1987/88-2004/05	SP	6794	0	10	0	1118	17	6271	92	6794	100	Incl. dwarf; 87/88-88/89. no microsats.
2005/06-2013/14	SP	3884	0	549	162	2645	68	3173	82	3884	100	Some missing by the 3/11 tsunami in 2011.
2014/15	SP	0	0	0	0	0	0	0	0	0	0	No catch.
Ant. fin whale												
2005/06-2013/14	SP	18	0	0	0	18	100	18	100	18	100	
2014/15	SP	0	0	0	0	0	0	0	0	0	0	No catch.
NP humpback whale												
2001-2014	BC	51	0	0	0	51	100	51	100	51	100	
2015	BC	10	0	0	0	10	100	10	100	10	100	
NP right whale												
2001-2014	BC	2	0	1	0	2	100	1	50	1	50	Missing by the 2011 tsunami, no microsats.
2015	BC	0	0	0	0	0	0	0	0	0	0	No BC.
NP fin whale												
2001-2014	BC	10	0	0	0	10	100	10	100	10	100	
2015	BC	0	0	0	0	0	0	0	0	0	0	No BC.

1	key to sample types: SP=special permit catch, C=commercial catch, BC=bycatch, ST=stranding.
2	number of whales that potentially entered by the previous years and enters (new year) the markets

Appendix 3

An update of the Norwegian minke whale DNA register

Hans J. Skaug

University of Bergen and Institute of Marine Research

footnote #	1	2	3	4	5	6	7	8	9	10	11	12
Species/Year	type	# whales	# duplicate	# missing	# lab problem	#mtDNA	%mtDNA	#msat	%msat	sex analyzed	% sexed	note
NP minke whale												
1997-2014	C	10061	101	65	2	9994	100	9994	100	9994	100	
2015	C	660	8	2	0	658	100	658	100	658	100	

1 key to sample types: SP=special permit catch, C=commercial catch, BC=bycatch, ST=stranding.
2 number of whales that potentially entered by the previous years and enters (new year) the markets
3 number of occurrences (tissues) sample switching on board the vessels as detected by comparison of genetic profiles.
4 number of individuals for which tissue samples are missing for other reasons than sample switching.
5 genetic laboratory not able to obtain microsatellite profiles mtDNA haplotypes from tissue samples.
6 number of samples analyzed for mitochondrial control region
7 % of total samples analyzed for mitochondrial control region
8 number of samples analyzed for microsatellites
9 % of total samples analyzed for microsatellites
10 number of samples analyzed for sex
11 % of total samples analyzed for sex
12 other problems or information

Appendix 4

STATUS OF THE ICELANDIC WHALE DNA REGISTER

Christophe Pampoulie and Gisli A. Vikingsson

Practical arrangements regarding the establishment of the Icelandic DNA register were concluded in 2007. The Marine Research Institute, Reykjavik, is responsible for the establishment and maintenance of the registry that is of the same format as the Norwegian DNA registry. An ORACLE database has now been created and contains all genotyped individuals information as well as tissue collected ID of individuals collected but not genotyped. In parallel, a DNA tissue bank has been achieved and is now fully functional.

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 (2003-07) and recent commercial catches (2008-15), as well as from commercial NA fin whale catches have been genotyped and information stored in the database.

footnote #	1	2	3	4	5	6	7	8	9	10	11	12
Species/Year	Type	# whales	# duplicate	# missing	# lab problem	#mtDNA	%mtDNA	#msat	%msat	sex analysed	% sexed	note
NA minke whale												
2003-2007	SP	189	0	0	0	189	100	189	100	189	100	
2008-2014	C	349	0	0	0	338	97	341	98	343	98	
2015	C	29	0	0	0	24	83	24	83	24	83	
NA fin whale												
2006-2014	C	534	0	0	0	534	100	534	100	534	100	
2015	C	154	0	0	0	154	100	154	100	154	100	

1 key to sample types: SP=special permit catch, C=commercial catch, BC=bycatch, ST=stranding.
2 number of whales that potentially entered by the previous years and enters (new year) the markets
3 number of occurrences (tissues) sample switching on board the vessels as detected by comparison of genetic profiles.
4 number of individuals for which tissue samples are missing for other reasons than sample switching.
5 genetic laboratory not able to obtain microsatellite profiles mtDNA haplotypes from tissue samples.
6 number of samples analysed for mitochondrial control region
7 % of total samples analysed for mitochondrial control region
8 number of samples analysed for microsatellites
9 % of total samples analysed for microsatellites
10 number of samples analysed for sex
11 % of total samples analysed for sex
12 other problems or information