

MOLECULAR ECOLOGY RESOURCES

Supplemental Information for:

A pan-cetacean MHC amplicon sequencing panel developed and evaluated in combination with genome assemblies

Dorothea Heimeier^{1,*}, Ellen C. Garland^{2,*}, Franca Eichenberger², Claire Garrigue^{3,4},
Adriana Vella⁵, C. Scott Baker⁶, , Emma L. Carroll^{1,*}

Table of Contents

SUPPORTING METHODS	2
PRIMER SEQUENCES AND PCR PROTOCOL	2
SUPPORTING TABLES	2
TABLE S1: OVERVIEW OF CETACEAN GENOME ASSEMBLIES USED IN THIS STUDY (SEE EXCEL FILE)	2
TABLE S2: TABLE WITH PRIMER SEQUENCES AND PCR PROTOCOL	2
TABLE S3: ORIGIN OF SAMPLES USED IN THIS STUDY INCLUDING COUNTRY OF ORIGIN, SOURCE AND PERMITS.....	3
TABLE S4: SUMMARY OF NUMBER OF INDIVIDUALS (N) SEQUENCED AND SUMMARY STATISTICS OF READ NUMBERS, MAPPING AND CLUSTERING.	4
TABLE S5: SUMMARY OF GENE SPECIFIC STATISTICS FOR SAMPLES FROM ODONTOCETES, MYSTICETES, HUMPBACK WHALES, AND SOUTHERN RIGHT WHALES.....	5
TABLE S6: SUMMARY OF AMPLICON PANEL CLASS I ALLELE RESULTS IN WHALES AND DOLPHINS.....	7
TABLE S7: MHC CLASS IIA ALLELES AT EACH LOCUS FOR EACH INDIVIDUAL SAMPLE.	8
TABLE S8: LIST OF CURATED ALLELES FROM CURRENT STUDY	9
SUPPLEMENTARY FIGURES	11
FIGURE S1: READ DEPTHS OF CLUSTERS FOR EACH CLASS IIA GENE FROM ALL SEQUENCING RUNS COMBINED.	11
<i>Mysticetes and Odontocetes</i>	11
<i>Humpback whales</i>	12
<i>Southern right whales</i>	13
FIGURE S2: READ DEPTHS OF CLUSTERS FOR EACH CLASS I GENE FROM ALL SEQUENCING RUNS COMBINED.	14
<i>Mysticetes and Odontocetes</i>	14
<i>Humpback whales</i>	14
<i>Southern right whales</i>	15
FIGURE S3: DEGREE OF CHANGE (DOC) CURVES FOR CLASS I AMPLICON READS FOR COMBINED RUNS.	16
<i>Mysticetes and Odontocetes</i>	16
<i>Humpback whales</i>	18
<i>Southern right whales</i>	20
FIGURE S4: DQB AND DRB ALIGNMENT OF GENES AT AMPLICON LENGTH FROM GENOMES.....	22
FIGURE S5: NEIGHBOR-JOINING PHYLOGENETIC TREES FOR AMPLICON LENGTH GENES FOR A) DRA, B) DRB, C) DQA, AND D) DQB.	24
ACKNOWLEDGEMENTS AND REFERENCES FOR GENOME ASSEMBLIES AS PROVIDED ON DNA-ZOO	28
<i>Mysticetes</i>	28

MOLECULAR ECOLOGY RESOURCES

<i>Odontocetes</i>	29
REFERENCES.....	34

Supporting Methods

Primer sequences and PCR protocol

Each PCR was carried out in a final volume of 20 µl with 3 µl template DNA at 50 ng genomic DNA, 4 µl 5X reaction buffer, 4 µl Phusion Plus GC enhancer, 0.2 µl PhusionPlus™ Polymerase (5 U/µl), 0.6 µl dNTPs and 0.8 µl each of forward and reverse primer. The finalized thermal cycling profile for the PCR was: 98°C for 30 seconds, 25 (DQA, DQB, DRA, DRB-a) or 30 cycles (class I) of 98°C for 5 seconds, 62°C for 10 seconds and 72°C for 30 seconds, followed by an extension step at 72°C for 5 minutes.

Supporting Tables

Table S1: Overview of cetacean genome assemblies used in this study (see excel file).

If assemblies from the same species were different from each other they are listed separately. Length of complete MHC region and class I and class IIa region between framework genes is given in kilobases (kb). Grey = at least one of the framework genes could not be identified. Red = an assembly was not usable for MHC plus explanation. Position of extracted genes are given from TRIM26 gene; grey = gene has not been identified in this assembly. If the gene was annotated in the assembly, the annotation reference is given without position information.

Table S2: Table with primer sequences and PCR protocol.

Each primer pair was contained the Illumina forward sequencing primer binding site (“TCGTCCGGCAGCGTCAGATGTGTATAAGAGACAG”) and a 5’ extension on the reverse primer containing the Illumina reverse sequencing primer site (“GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG”).

Gene	Primer -forward 5’-3’	Length in bp	Primer -reverse 5’-3’	Length in bp	Predicted length of amplicon excl primer
DRA	TCTCTCCTGGTCCCACCCTA	21	CCTGGGAGGGAAAGGCAGT	19	349bp
DRB-a	GAGCCCGCCAGGTRATCAG	20	CCCGAACGCAGTCTCCTCTG	20	389bp
DQA	ACATGGTTCTTCTCTCYCTCTG	24	GGGAACAAGAGAGTGAGGYCTG	22	334bp
DQB	GTTGAGCGGCGGGTTTCAG	19	TCTCGGCAGGKGRGGGTG	18	384bp
Class I	AGGCTCCCACTCCCTGAGG	19	CGGGTCYSGCAATGTTGGG	19	341bp

MOLECULAR ECOLOGY RESOURCES

Table S3: Origin of samples used in this study including country of origin, source and permits.

Common name	Species name	four letter code	Number of Samples	Sample codes	Source of samples	Country of origin	Permit
Southern right whale	<i>Eubalaena australis</i>	<i>Euuu</i>	30	see FigureS3C	biopsies	NZ	1)
Pygmy blue whale	<i>Balaenoptera musculus brevicauda</i>	<i>Bamu</i>	1	Bamu-U18-023	stranding	NZ	NZCeTA
Eden's whale	<i>Balaenoptera edeni</i>	<i>Baed</i>	2	Bed31, Bed34	strandings	NZ	NZCeTA
Common minke whale	<i>Balaenoptera acutorostrata</i>	<i>Baac</i>	2	Bac16; Bac-QLM-11-91	strandings	NZ	NZCeTA
Humpback whale	<i>Megaptera novaeangliae</i>	<i>Meno</i>	25	see FigureS3B	biopsies	New Caledonia	2)
		<i>Meno</i>	1	Meno-NZ12A	stranding	NZ	NZCeTA
Bottlenose dolphin	<i>Tursiops truncatus</i>	<i>Tutr</i>	2	Ttr05-BOI106, Ttr-JB04-05	biopsies	NZ	3)
		<i>Tutr</i>	1	Tutr-U18-059	strandings	NZ	NZCeTA
Hector's dolphin	<i>Cephalorhynchus hectori</i>	<i>Cehec</i>	2	Che11CB124, Che12CB043	biopsies	NZ	4)
Long-finned pilot whale	<i>Globicephala melas</i>	<i>Glme</i>	2	Glo093, Glo099	strandings	NZ	NZCeTA
Dusky dolphin	<i>Lagenorhynchus obscurus</i>	<i>Laob</i>	2	Lob06, Lob07	strandings	NZ	NZCeTA
Risso's dolphin	<i>Grampus griseus</i>	<i>Grgr</i>	1	Grgr-M1	stranding	Malta	8)
		<i>Grgr</i>	1	Grgr-U19-016	stranding	NZ	NZCeTA
Rough-toothed dolphin	<i>Steno bredanensis</i>	<i>Sbr</i>	2	Sbr-04-FP03; Sbr-07-SA03	biopsies	French Polynesia; Samoa	5)
Short-beaked common dolphin	<i>Delphinus delphis</i>	<i>Dede</i>	1	Dde-M1	stranding	Malta	8)
Striped dolphin	<i>Stenella coeruleoalba</i>	<i>Stco</i>	1	Sco-M1	stranding	Malta	8)
Dwarf sperm whale	<i>Kogia sima</i>	<i>Kosi</i>	1	Ksi-NC05-111	stranding	New Caledonia	6)
Pygmy sperm whale	<i>Kogia breviceps</i>	<i>Kobr</i>	1	Kbr106	stranding	NZ	NZCeTA
Blainville's beaked whale	<i>Mesoplodon densirostris</i>	<i>Mede</i>	2	Mede1, Mede2	biopsies	French Polynesia -Raitea	5)
Cuvier's beaked whale	<i>Ziphius cavirostris</i>	<i>Zica</i>	2	Zca-NZ12-99; Zica-U19-001	strandings	NZ	NZCeTA
Gervais' beaked whale	<i>Mesoplodon europaeus</i>	<i>Meeu</i>	1	Meu-SW7444	strandings	USA	7)
Gray's beaked whale	<i>Mesoplodon grayi</i>	<i>Megr</i>	2	Megr-U17-093; Megr-U18-013	strandings	NZ	NZCeTA

1) Field work was conducted under University of Auckland Animal Ethics approved protocol 002072 to Emma Carroll and under New Zealand Marine Mammal Protection Act Permit 84845-MAR and Marine Reserve Act Permit 87513-MAR.

2) Biopsy samples were collected under permits to Claire Garrigue. No permit requirements pre 2004; Provence Sud -2006: 60241766/DRN/ENV; 2007: 60241115/DENV/MT/DP; 2009: 1462.2008/PS; 2010: 383-2010/ARR/DENV; and 2011: 3313-2010/ARR/DENV.

3) Biopsy samples were collected under the permit to C.S.B. from the New Zealand Department of Conservation and animal ethics protocols AEC/02/2002/R9 and AEC/02/2005/R334 from the University of Auckland

4) Biopsy samples were collected under permit Rnw/HO/2009/03 from DOC issued to CSB, according to Animal Care and Use Protocol 3805 approved by the Oregon State University Institutional Animal Care and Use Committee, and University of Auckland Animal Ethics Protocol AEC/02/2008/ R658.

5) Rough-toothed dolphin samples were collected in collaboration with Scott Baker and Marc Oremus from French Polynesia with support of Pew Marine Conservation Fellowship, as part of the project 'A pattern of dolphins'. Research permit was delivered by the Direction de l'Environnement of FP to Michael Poole's marine mammal research program and from Samoa by Juney Ward as part of the Ministry of Environment.

6) Under stranding permit issue number 6021-1028/DRN/ENV.

7) SWFSC Marine Mammal and Sea Turtle Research (MMaSTR) collection

8) Malta cetacean samples came from the collection by the Conservation Biology Research Group, Department of Biology, University of Malta under Adriana Vella's care and lead.

Environment and Research Authority (ERA) Permits were issued in Adriana Vella's name. These samples were forwarded by the University of Malta under both CITES and EPD permits in July 1999.

MOLECULAR ECOLOGY RESOURCES

Table S4: Summary of number of individuals (N) sequenced and summary statistics of read numbers, mapping and clustering.

Mysticetes* excluding humpback and southern right whales. Odontocetes = toothed whales; Mysticetes = baleen whales.

	N	N <1k reads	# total reads	# avg reads/ind	# reads mapped	% of total	class I					class IIa				
							# reads mapped	% of total	# avg reads/ind	# reads clustered	% of mapped	# reads mapped	% of total	# avg reads/ind	# reads clustered	% of mapped
Odontocetes	24	0	660,698	27,529	631,233	96	405,216	61	16,884	315,602	78	226,017	34	9,417	200,124	89
Mysticetes*	5	2	120,631	24,126	119,346	99	63,989	53	12,798	51,356	80	55,357	46	11,071	47,892	87
Humpback whales	26	0	778,390	29,938	777,185	100	499,472	64	19,210	395,530	79	277,713	36	10,681	248,322	89
Southern right whales	30	1	285,089	9,503	283,468	99	183,215	64	6,107	158,748	87	100,253	35	3,342	87,125	87
Total	85		1,844,808	21,704	1,811,232	98	1,151,892	62	13,552	921,236	80	659,340	36	7,757	583,463	88

	class II DRA				class II DRB				class II DQA				class II DQB			
	# reads mapped	% of total	# reads clustered	% of mapped	# reads mapped	% of total	# reads clustered	% of mapped	# reads mapped	% of total	# reads clustered	% of mapped	# reads mapped	% of total	# reads clustered	% of mapped
Odontocetes	27,161	4	22,920	84	114,909	17	103,622	90	43,249	7	37,680	87	40,698	6	35,902	88
Mysticetes*	8,034	7	6,913	86	23,004	19	20,468	89	3,810	3	3,434	90	20,509	17	17,077	83
Humpback whales	29,114	4	25,815	89	65,221	8	58,657	90	55,572	7	50,872	92	127,806	16	112,978	88
Southern right whales	21,862	8	18,036	82	18,308	6	15,557	85	34,225	12	31,300	91	25,858	9	22,232	86
Total	86,171	5	73,684	86	221,442	12	198,304	90	136,856	7	123,286	90	214,871	12	188,189	88

MOLECULAR ECOLOGY RESOURCES

Table S5: Summary of gene specific statistics for samples from odontocetes, mysticetes, humpback whales, and southern right whales.

Statistics include number of individuals that amplified a gene (N), average reads per allele and individual (avg read/all/ind), total number of alleles (# all) and their pairwise identity in % (pw id%), and number of alleles identical to previously described alleles from assemblies and/or Genbank (assbl/GB).

	DRA (351bp)						DRB-a (405bp)					
	Total N	N	avg read/all/ind (+SD)	# all	Pw id %	alleles prev described (assbl/ GB)	N	avg read/all/ind (+SD)	# all	Pw id %	alleles prev described (assbl/ GB)	
Odontocetes	24	24	640 (+800)	20	97.5	5/6	22	2,731 (+2470)	27	93.5	0/1	
Mysticetes*	5	5	688 (+1262)	5	99.6	4/2	5	2,791 (+4020)	4	97.7	0/1	
Humpback whales	26	26	852 (+636)	1	na	1/0	26	1,352 (+1060)	12	98.0	0/5	
Southern right whales	30	30	455 (+434)	2	98.0	1/0	28	298 (+724)	6	98.7	1/3	
Total	85	85	659	28	96.3	15	81	1,793	49	92.2	11	

	DQA (336bp)						DQB (385bp)					
	Total N	N	avg read/all/ind (+SD)	# all	Pw id %	alleles prev described (assbl/ GB)	N	avg read/all/ind (+SD)	# all	Pw id %	alleles prev described (assbl/ GB)	
Odontocetes	24	23	952 (+567)	26	96.0	3/7	21	908 (+992)	29	93.3	1/10	
Mysticetes*	5	3	750 (+1366)	3	98.6	3/1	3	2,901 (+2000)	4	96.1	1/1	
Humpback whales	26	26	1,606 (+1407)	1	na	1/0	26	2,109 (+1204)	14	97.4	1/5	
Southern right whales	30	30	788 (+566)	2	99.7	1/1	29	365 (+1150)	15	97.4	0	
Total	85	82	1024	32	95.1	13	79	1,571	62	93.6	19	

MOLECULAR ECOLOGY RESOURCES

	Class I (318bp/341bp)					
	Total N	N	avg read/all/ ind (+-SD)	# all	Pw id %	alleles prev described (assbl/ GB)
Odontocetes	24	22	2,307 (+-2256)	105	92.7	1/5
<u>Mysticetes*</u>	5	2	2,432 (+-2395)	13	93.6	2/0
Humpback whales	26	25	1,693 (+-1137)	33	92.2	1/0
Southern right whales	30	26	962 (+-959)	32	93.3	3/0
Total	85	75	1,849	183	92.2	12

MOLECULAR ECOLOGY RESOURCES

Table S6: Summary of amplicon panel class I allele results in whales and dolphins.

Sample indicates the sample code of the sample used, except for humpback whale (Meno) and southern right whale (Euau) where the species code and sample size (n) are given. Shown is the number of class I alleles (# alleles) identified from 72 samples from 12 different cetacean species mapped the cetacean genome assembly indicated by “genome” at 95% and 92% overlap identity. An exact match of one allele to the genomic sequence in the assembly is in green shading and bold numbers. When a sequence differed by only 1-3 bp the shading is in green without bold numbers.

Sample	cetacean family	# alleles	genome	number of alleles mapped at 95% (92%) to genome					un-mapped
				κ class I a	κ class I b	κ class I c	middle class I	β class I	
Bamu-U18-023	Balaenopteridae	7	Blue whale	1 (3)			1 (2)	1 (2)	4 (0)
Bed-34-1	Balaenopteridae	6	Rice's whale		1 (1)		2 (4)	1 (1)	2 (0)
Meno (n=26)	Balaenopteridae	33	Blue whale	3 (12)	3 (5)		0 (6)	1 (9)	26 (1)
Euau (n=30)	Balaenidae	32	Southern right whale		2 (2)		12 (25)	2 (5)	16 (0)
Che12CB043	Delphinidae	3	Bottlenose dolphin				0 (2)	0 (1)	3 (0)
Dde-M1	Delphinidae	7	Bottlenose dolphin			0 (1)	1 (4)	0 (2)	6 (0)
Glo93	Delphinidae	8	Bottlenose dolphin		2 (2)	1 (3)	3 (3)		2 (0)
Glo99	Delphinidae	3	Bottlenose dolphin		1 (1)		2 (2)		
Ggr-M1	Delphinidae	7	Bottlenose dolphin			3	3 (3)	1 (1)	
Ggr-U19-016	Delphinidae	5	Bottlenose dolphin			1 (1)	2 (2)	2 (2)	
Lob06	Delphinidae	8	Bottlenose dolphin			3 (3)	1 (3)	0 (2)	4 (0)
Lob07	Delphinidae	7	Bottlenose dolphin		0 (1)	1 (2)	1 (2)	2 (2)	3 (0)
Sbr-04-FP03	Delphinidae	7	Bottlenose dolphin				2 (2)	3 (4)	2 (0)
Sbr-07-SA03	Delphinidae	7	Bottlenose dolphin		1 (1)	0 (2)	2 (2)	2 (2)	2 (0)
Sco-M1	Delphinidae	5	Bottlenose dolphin		1 (1)	0 (1)	0 (1)	1 (2)	3 (0)
Ttr05-BOI106	Delphinidae	9	Bottlenose dolphin		1 (1)	1 (1)	2 (3)	3 (4)	2 (0)
Ttr-JB04-05	Delphinidae	5	Bottlenose dolphin			1 (2)	1 (1)	1 (2)	2 (0)
Ttr-U18-059	Delphinidae	10	Bottlenose dolphin		2 (2)		2 (4)	2 (4)	4 (0)

MOLECULAR ECOLOGY RESOURCES

Table S7: MHC class IIa alleles at each locus for each individual sample.

Light red shading indicates allele was called with less than 20 reads. Green writing indicates allele is already present on GenBank (see also Table S8).

		DRA		DRB-a		DQA		DQB	
	<20 reads	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2	Allele 1	Allele 2
Balenopteridae	Bac-16	Baac-DRA**01	Baac-DRA**02	Baac-DRB-a**01					
Balenopteridae	Bac-QLM	Baac-DRA**01	Baac-DRA**02	Baac-DRB-a**01					
Balenopteridae	Bamu-U18-0236	Bamu-DRA**01		Bamu-DRB-a**01		Bamu-DQA**01	Bamu-DQA**02	Bamu-DQB**01	Bamu-DQB**02
Balenopteridae	Bed-31-1	Baed-DRA**02		Baed-DRB-a**01		Baed-DQA**01		Baed-DQB**02	
Balenopteridae	Bed-34-1	Baed-DRA**01	Baed-DRA**02	Baed-DRB-a**01	Baed-DRB-a**02	Baed-DQA**01		Baed-DQB**02	Baed-DQB**01
Delphinidae	Cehc11-CB124	Cehc-DRA**01		Cehc-DRB-a**01		Cehc-DQA**01		Cehc-DQB**01	
Delphinidae	Che12CB043	Cehc-DRA**01	Cehc-DRA**02	Cehc-DRB-a**01		Cehc-DQA**02	Cehc-DQA**01	Cehc-DQB**01	Cehc-DQB**02
Delphinidae	Dde-M1	Dede-DRA**01	Dede-DRA**02	Dede-DRB-a**01	Dede-DRB-a**02	Dede-DQA**01	Dede-DQA**02	Dede-DQB**01	Dede-DQB**02
Delphinidae	Glo93	Glme-DRA**01	Glme-DRA**02	Glme-DRB-a**01		Glme-DQA**01			
Delphinidae	Glo99	Glme-DRA**01		Glme-DRB-a**02	Glme-DRB-a**03	Glme-DQA**01		Glme-DQB**01	
Delphinidae	Grgr-M1			Grgr-DRB-a**01	Grgr-DRB-a**02	Grgr-DQA**01	Grgr-DQA**02	Grgr-DQB**01	Grgr-DQB**02
Delphinidae	Grgr-U19-016	Grgr-DRA**01		Grgr-DRB-a**03		Grgr-DQA**01	Grgr-DQA**02	Grgr-DQB**03	Grgr-DQB**04
Kogiidae	Kbr-106	Kobr-DRA**01				Kobr-DQA**01	Kobr-DQA**02		
Kogiidae	Ksi-NC05-111	Kosi-DRA**01		Kosi-DRB-a**01		Kosi-DQA**01	Kosi-DQA**02	Kosi-DQB**01	Kosi-DQB**02
Delphinidae	Lob06	Laob-DRA**01		Laob-DRB-a**01		Laob-DQA**01	Laob-DQA**02	Laob-DQB**01	
Delphinidae	Lob07	Laob-DRA**01		Laob-DRB-a**01		Laob-DQA**01		Laob-DQB**01	
Ziphiidae	Mede1	Mede-DRA**01		Mede-DRB-a**01		Mede-DQA**01		Mede-DQB**01	Mede-DQB**02
Ziphiidae	Mede2	Mede-DRA**01	Mede-DRA**02	Mede-DRB-a**01	Mede-DRB-a**02	Mede-DQA**01	Mede-DQA**02	Mede-DQB**03	Mede-DQB**04
Ziphiidae	Meu-SW7444	Meeu^01		Meeu-DRB-a**02	Meeu-DRB-a**03	Meeu-DQA**01	Meeu-DQA**02		
Ziphiidae	Megr-U17-093	Megr-DRA**01		Megr-DRB-a**01		Megr-DQA**01		Megr-DQB**01	Megr-DQB**03
Ziphiidae	Megr-U18-013	Megr-DRA**01		Megr-DRB-a**02	Megr-DRB-a**03	Megr-DQA**01	Megr-DQA**02	Megr-DQB**02	
Delphinidae	Sbr-04-FP03	Stbr-DRA**01		Stbr-DRB-a**01		Stbr-DQA**01		Stbr-DQB**01	Stbr-DQB**02
Delphinidae	Sbr-07-SA03	Stbr-DRA**01		Stbr-DRB-a**02	Stbr-DRB-a**03	Stbr-DQA**01		Stbr-DQB**03	Stbr-DQB**02
Delphinidae	Sco-M1	Stco-DRA**01				Stco-DQA**01		Stco-DQB**01	
Delphinidae	Ttr05-BO1106	Tutr-DRA**01		Tutr-DRB-a**03		Tutr-DQA**01	Tutr-DQA**02	Tutr-DQB**01	
Delphinidae	Ttr-JB04-05	Tutr-DRA**01		Tutr-DRB-a**03		Tutr-DQA**03		Tutr-DQB**04	
Delphinidae	Tutr-U18-059	Tutr-DRA**02	Tutr-DRA**03	Tutr-DRB-a**01	Tutr-DRB-a**02			Tutr-DQB**02	Tutr-DQB**03
Ziphiidae	Zca-NZ12-99	Zica-DRA**01		Zica-DRB-a**01	Zica-DRB-a**02	Zica-DQA**01		Zica-DQB**01	
Ziphiidae	Zica-U19-001	Zica-DRA**01		Zica-DRB-a**03	Zica-DRB-a**01	Zica-DQA**01	Zica-DQA**02	Zica-DQB**01	Zica-DQB**02

MOLECULAR ECOLOGY RESOURCES

Table S8: List of curated alleles from current study

Curated alleles from the current study and how they correspond to previously described alleles on GenBank or in published genome assemblies. Light red shading indicates allele was called with less than 20 reads. Yellow shading indicates alleles include stop codons in reading frame and predicted to be non-functional. Accession numbers of GenBank sequences are given, accession numbers for genome assemblies can be found in TableS1. Green text highlights alleles which sequences were found on GenBank or genome assemblies.

DRA	same allele as allele	same allele as in genome assembly	GenBank
Baac**01	Meno, Bamu	Meno, Bamu, Baac	
Baac**02			
Baed**01	Babo	Bari	Baom-DRA*01_EF375605
Baed**02			
Bamu**01	Meno, Baac	Meno, Bamu, Baac	
Cehc**01			Deca-DRA*0101_FM986350
Cehc**02			
Dede**01			
Dede**02			Deca-DRA*01_EF375603, Stco-DRA*04_EF375588
Euau**01		Euau, Eugl	
Euau**02			
Glme**01	Grgr	Glme	Grgr-DRA*0101_FM986351
Glme**02			
Grgr**01	Glme	Glme	Grgr-DRA*0101_FM986351
Kobr**01			
Kosi**01			
Mede**01	Megr, Meeu		
Mede**02			
Meeu**01	Mede, Megr		
Megr**01	Meeu, Mede		
Meno**01	Baac, Bamu	Meno, Bamu, Baac	
Laob**01		Saob	
Stbr**01		Stbr	
Stco**01			
Tutr**01		Tutr	Deca-DRA*0101_FM986350
Tutr**02			
Tutr**03			
Zica**01			

DRB	same allele as in genome assembly	GenBank
Baac**01		
Baed**01		
Baed**02		
Bamu**01		Bamu94001-DRB*3c_DQ354666
Cehc**01		
Dede**01		
Dede**02		
Euau**01		
Euau**02		
Euau**03	Euau	
Euau**04		EuauWA9511-DRB*1c_DQ354677
Euau**05		EuauWA9511-DRB*1c_DQ354677
Euau**06		EuauWA9511-DRB*1c_DQ354677
Glme**01		
Glme**02		
Glme**03		
Grgr**01		
Grgr**02		
Grgr**03		
Kosi**01		
Mede**01		
Mede**02		
Meeu**02		
Meeu**03		
Megr**01		
Megr**02		
Megr**03		
Meno**01		MenoSEA88060-DRB*4c_DQ354685
Meno**02		
Meno**03		MenoCa91037-DRB*6_DQ354681, MenoGM90012-DRB*7_DQ354682
Meno**04		
Meno**05		MenoWA94020-DRB*5_DQ354687
Meno**06		
Meno**07		
Meno**08		MenoWA94016-DRB*3_DQ354686
Meno**09		MenoSEA88060-DRB*2c_DQ354684
Meno**10		
Meno**11		
Meno**12		
Laob**01		CeheBP18-DRB_DQ354675
Stbr**01		
Stbr**02		
Stbr**03		
Tutr**01		(Tutr-DRB*04_EF507865, Tutr-DRB*06_EF507867, Tutr-DRB*05_EF507866) exon2 only
Tutr**02		
Tutr**03		(Tuad-DRB_EF017818, Tuad-DRB*07_EF507874, Tuad-DRB*05_EF507872) exon 2 only
Zica**01		
Zica**02		
Zica**03		

MOLECULAR ECOLOGY RESOURCES

DQA	same allele as in genome assembly	GenBank
Baed**01	Bari	
Bamu**01		
Bamu**02	Bamu	BamuCA9402-DQA_EU024800
Cehc**01		Cehe-DQA*02_EU024797
Cehc**02		Cehe-DQA*04_MH484168
Dede**01		Sogu-DQA4_FJ848540
Dede**02		
Euau**01	Eugl, Euau	Bamy9201-DQA_EU024801
Euau**02		
Glme**01		
Grgr**01		
Grgr**02	Stbr, Glme	GlmeNZ207-DQA_EU024803, Sogu-DQA3_FJ848539
Kobr**01		
Kobr**02		
Kosi**01		
Kosi**02		
Mede**01		
Mede**02		
Meeu**01		
Meeu**02		
Megr**01		
Megr**02		
Meno**01	Meno	
Laob**01		
Laob**02		Cehe-DQA*03_MH484167 GlmeNZ207-DQA_EU024803, Sogu-DQA3_FJ848539
Stbr**01	Stbr, Glme	
Stco**01		
Tutr**01		
Tutr**02		TutrNZ17-DQA_EU024798
Tutr**03	Soch	
Zica**01		
Zica**02		

DQB	same allele as in genome assembly	GenBank
Baed**01		
Baed**02		
Bamu**01	Bamu	
Bamu**02		BamuGC-DQB*18_KJ179635
Cehc**01		CeheNI01-DQB*2_DQ354629
Cehc**02		Cehe-DQB*05_MH484170
Dede**01		Cehe-DQB*06_MH484171
Dede**02		Stco-DQB*16_KY119141
Euau**01		
Euau**02		
Euau**03		
Euau**04		
Euau**05		
Euau**06		
Euau**07		
Euau**08		
Euau**09		
Euau**10		
Euau**11		
Euau**12		
Euau**13		
Euau**14		
Euau**15		
Glme**01		
Grgr**01		
Grgr**02		
Grgr**03		
Grgr**04		
Kosi**01		
Kosi**02		
Mede**01		
Mede**02		
Mede**03		
Mede**04		
Megr**01		
Megr**02		
Megr**03		
Meno**01		MenoGB0003-DQB*12c_DQ354
Meno**02		MenoSEA-DQB*21c_DQ354662
Meno**03	Meno	
Meno**04		MenoCA-DQB*6c_DQ354647
Meno**05		MenoSEA-DQB*21c_DQ354662
Meno**06		MenoCA-DQB*5c_DQ354646
Meno**07		
Meno**08		
Meno**09		
Meno**10		
Meno**11		
Meno**12		
Meno**13		
Meno**14		
Laob**01		
Stbr**01		CeheNI01-DQB*2_DQ354629
Stbr**02	Stbr	
Stbr**03		
Stco**01		Stco-DQB*17_KY119145
Tutr**01		Tutr-DQB*07_AB302050
Tutr**02		Tutr-DQB*10_AB302053
Tutr**03		Tutr-DQB*12_AB302055
Tutr**04		Tutr-DQB*09_AB302052
Zica**01		
Zica**02		

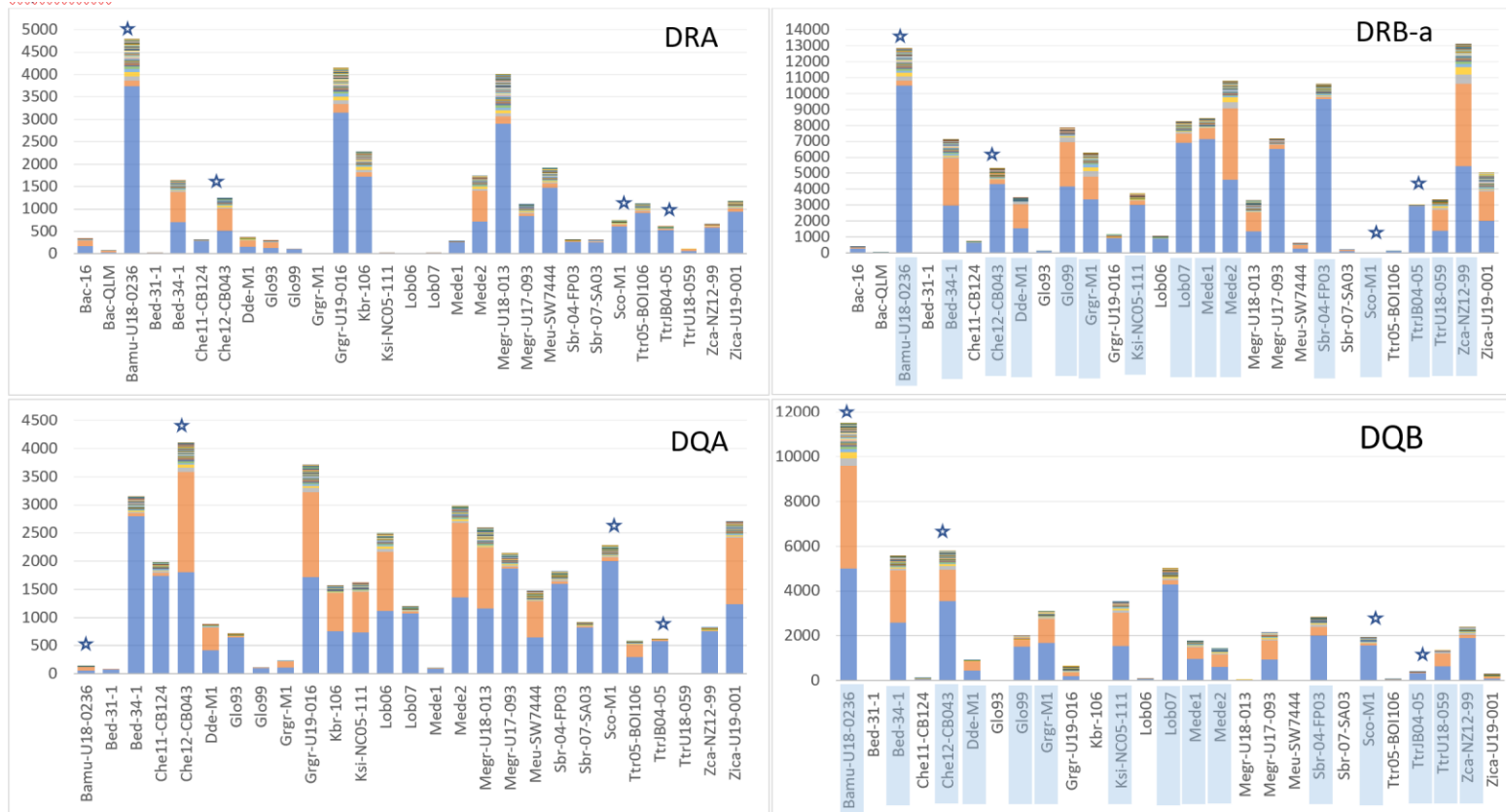
MOLECULAR ECOLOGY RESOURCES

Supplementary Figures

Figure S1: Read depths of clusters for each class IIa gene from all sequencing runs combined.

Samples with a star were technical repeat samples on the same plate. Samples and amplicons in blue shading were additionally sequenced on the MiSeq.

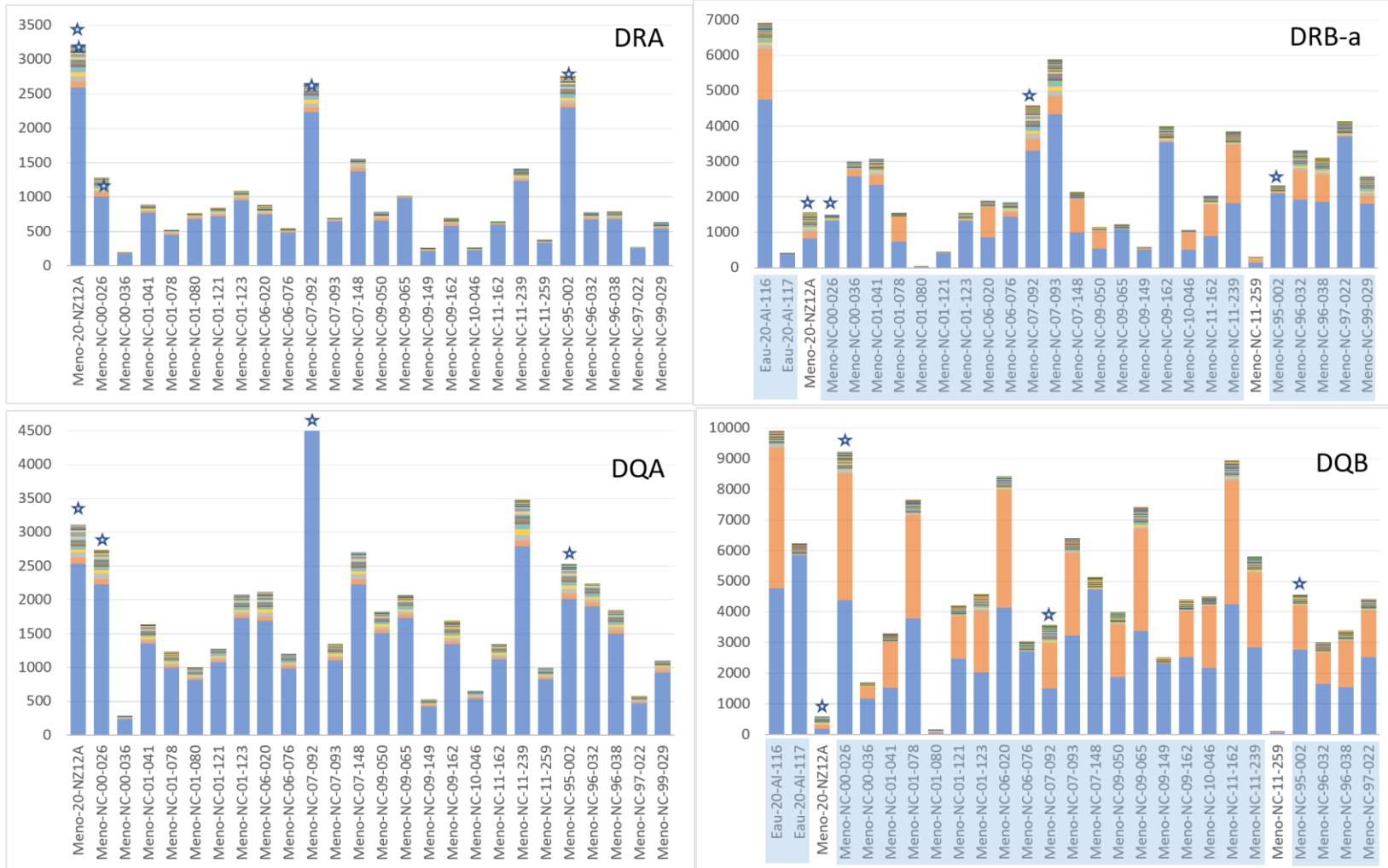
Mysticetes and Odontocetes



DRB-b

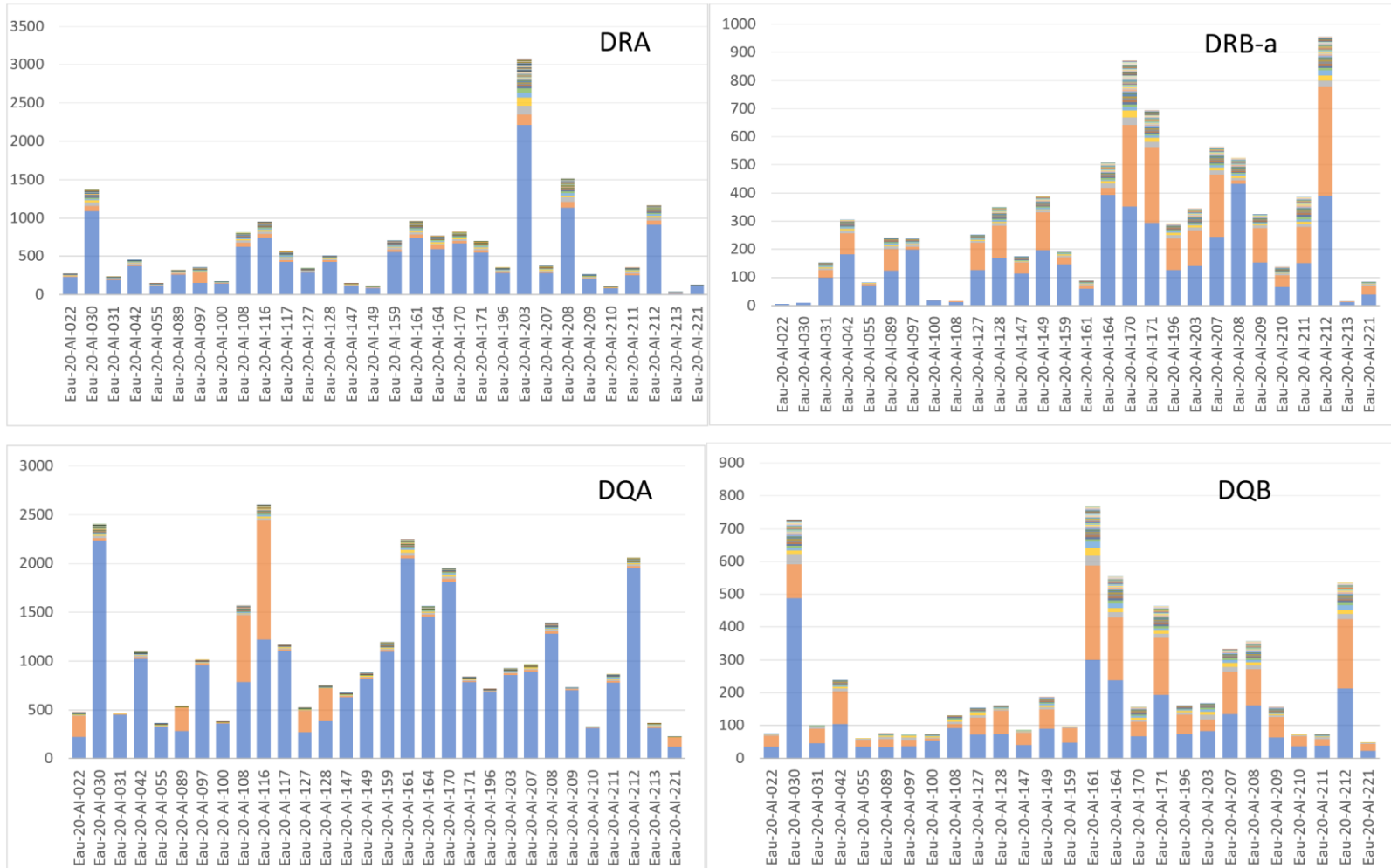
MOLECULAR ECOLOGY RESOURCES

Humpback whales (including 2 Southern right whales for which read depth was too high to fit in Southern right whale graph)



MOLECULAR ECOLOGY RESOURCES

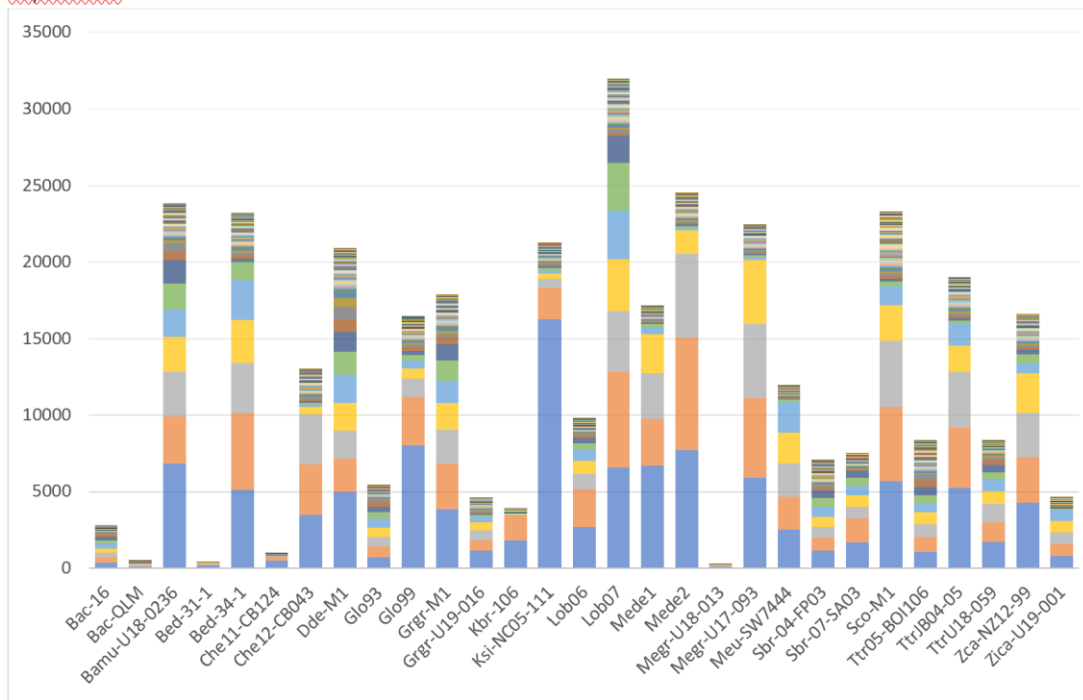
Southern right whales



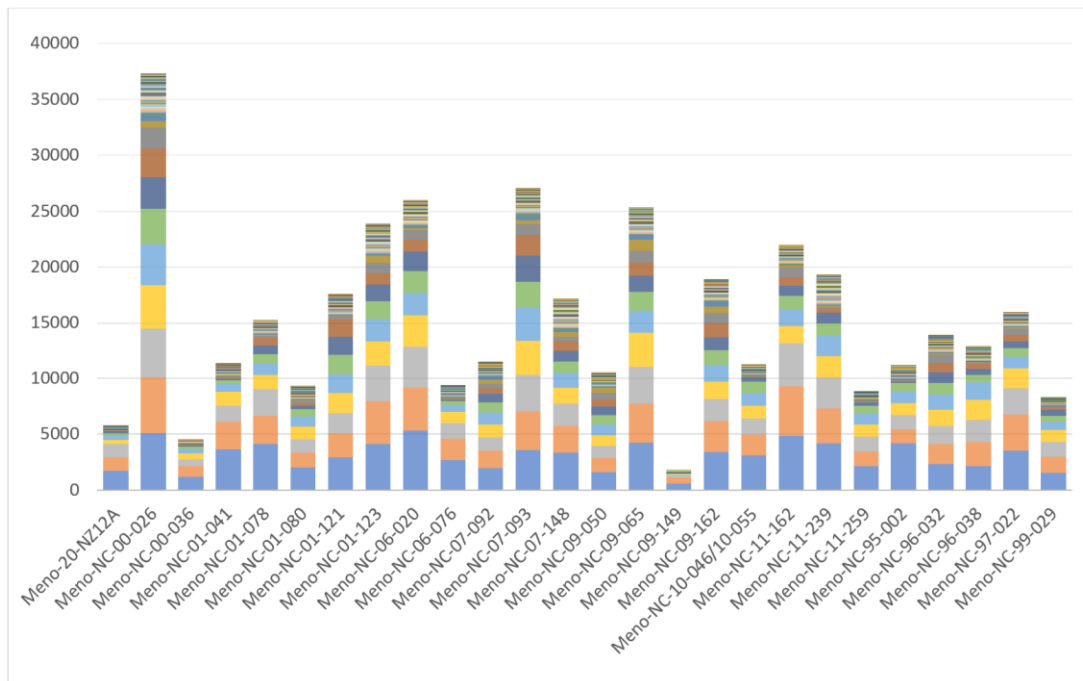
MOLECULAR ECOLOGY RESOURCES

Figure S2: Read depths of clusters for each class I gene from all sequencing runs combined.

Mysticetes and Odontocetes

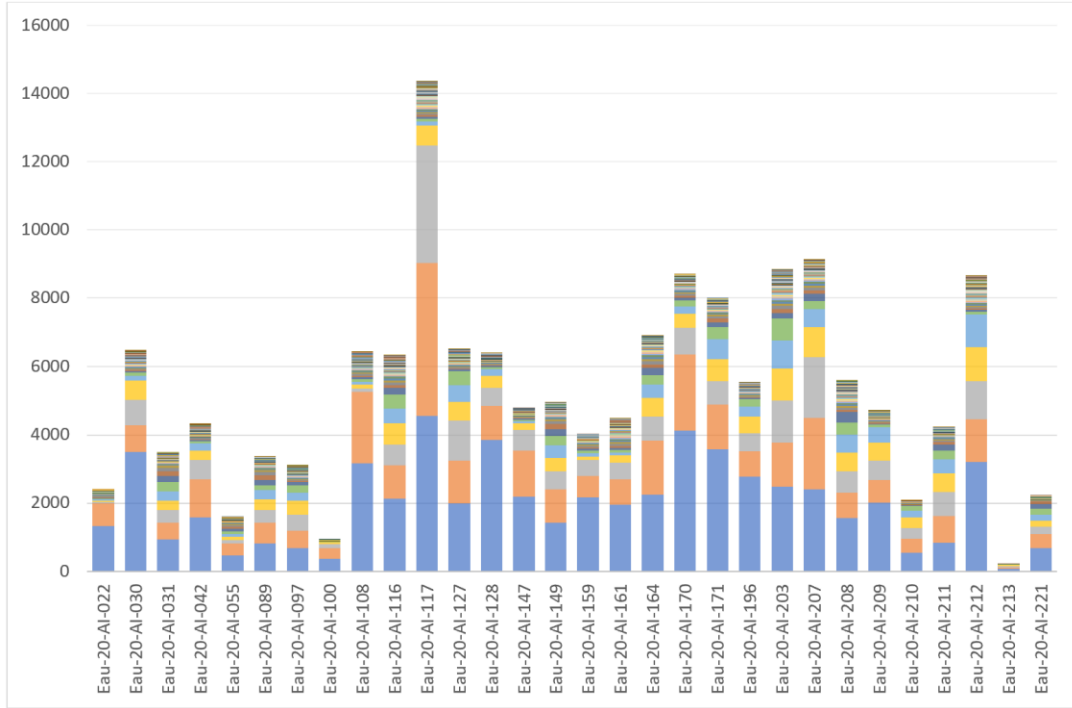


Humpback whales



MOLECULAR ECOLOGY RESOURCES

Southern right whales

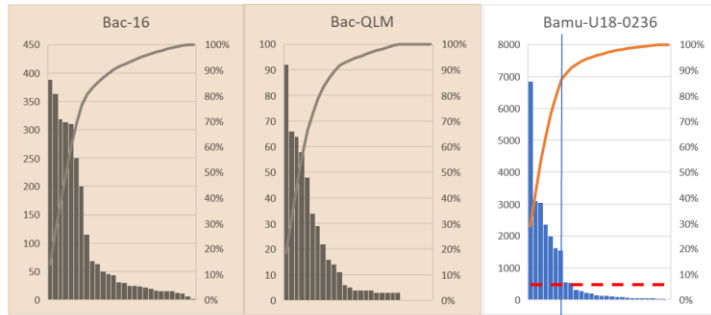


MOLECULAR ECOLOGY RESOURCES

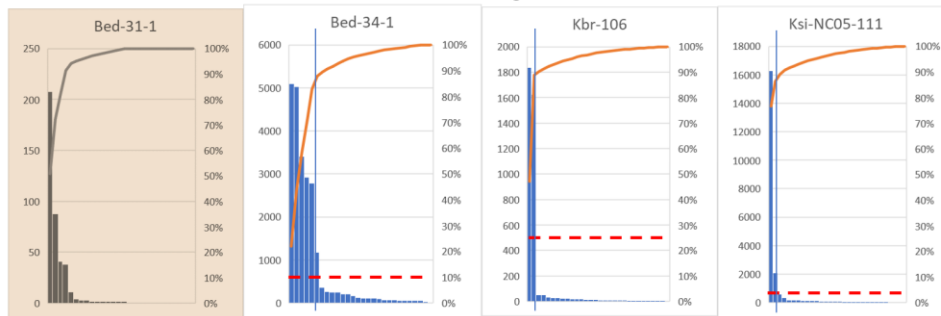
Figure S3: Degree of change (DoC) curves for class I amplicon reads for combined runs. Graphs shaded in red were excluded from further analysis due to low overall read numbers. The dashed red line indicates 500 reads.

Mysticetes and Odontocetes

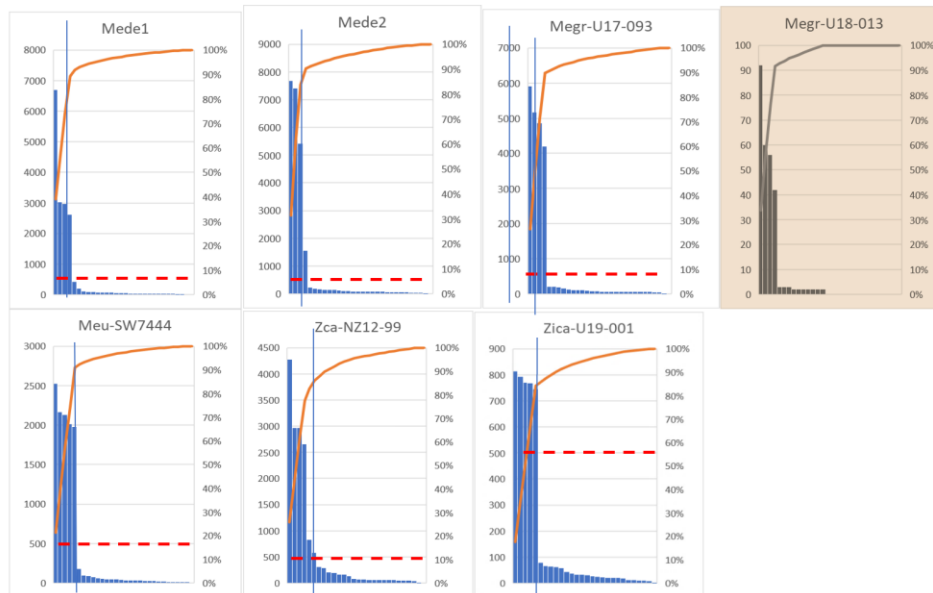
Balenopteridae



Kogiidae

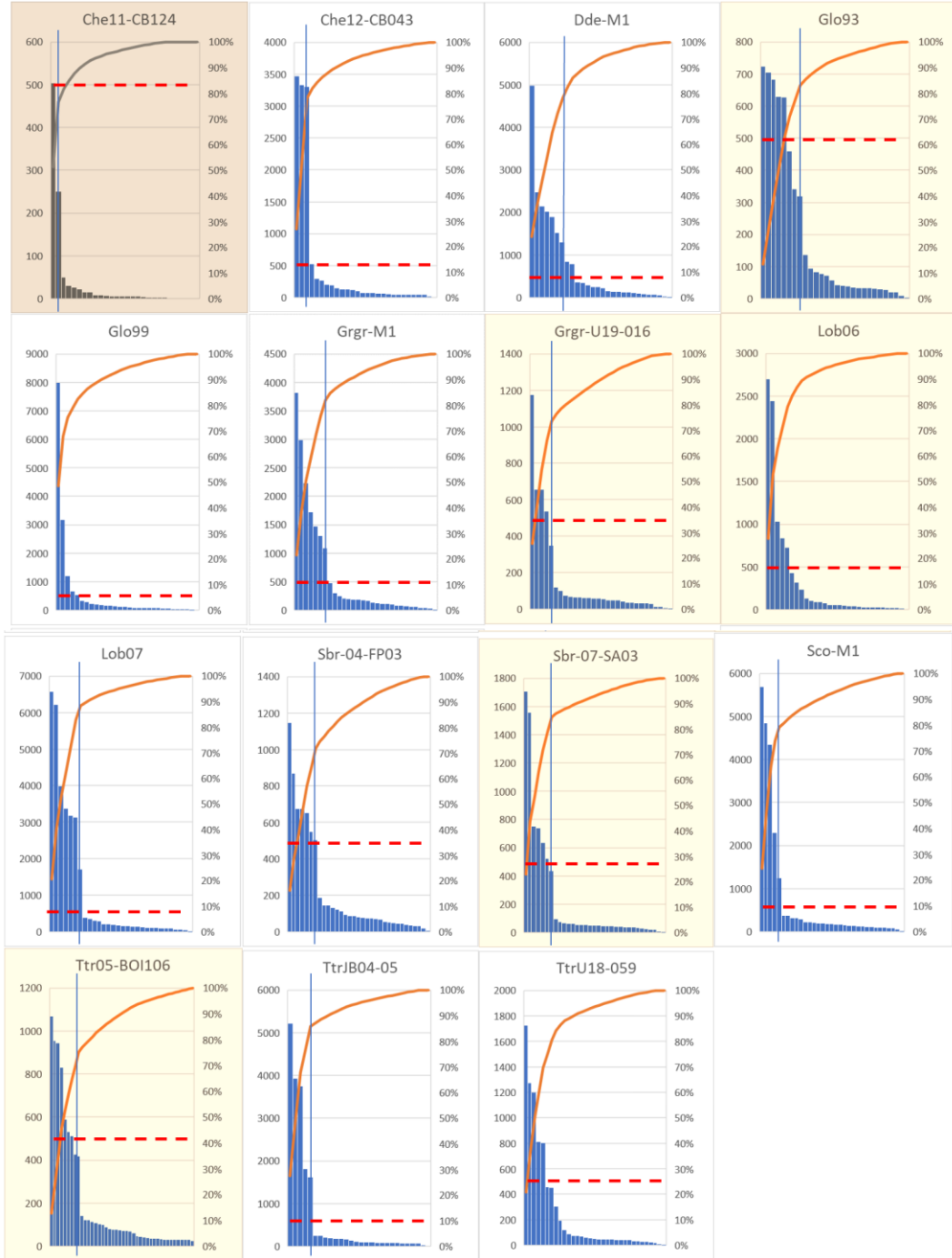


Ziphiidae



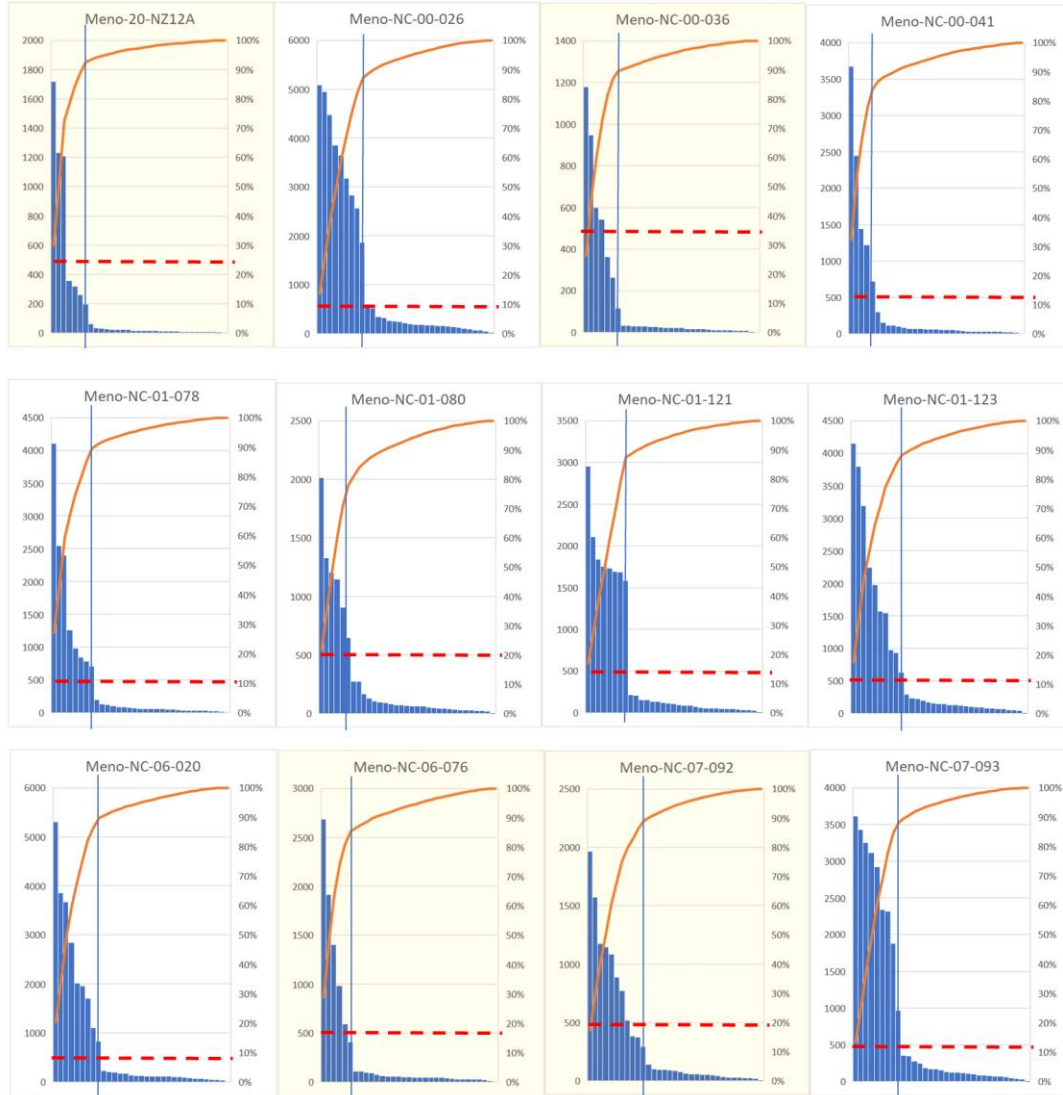
MOLECULAR ECOLOGY RESOURCES

Delphinidae

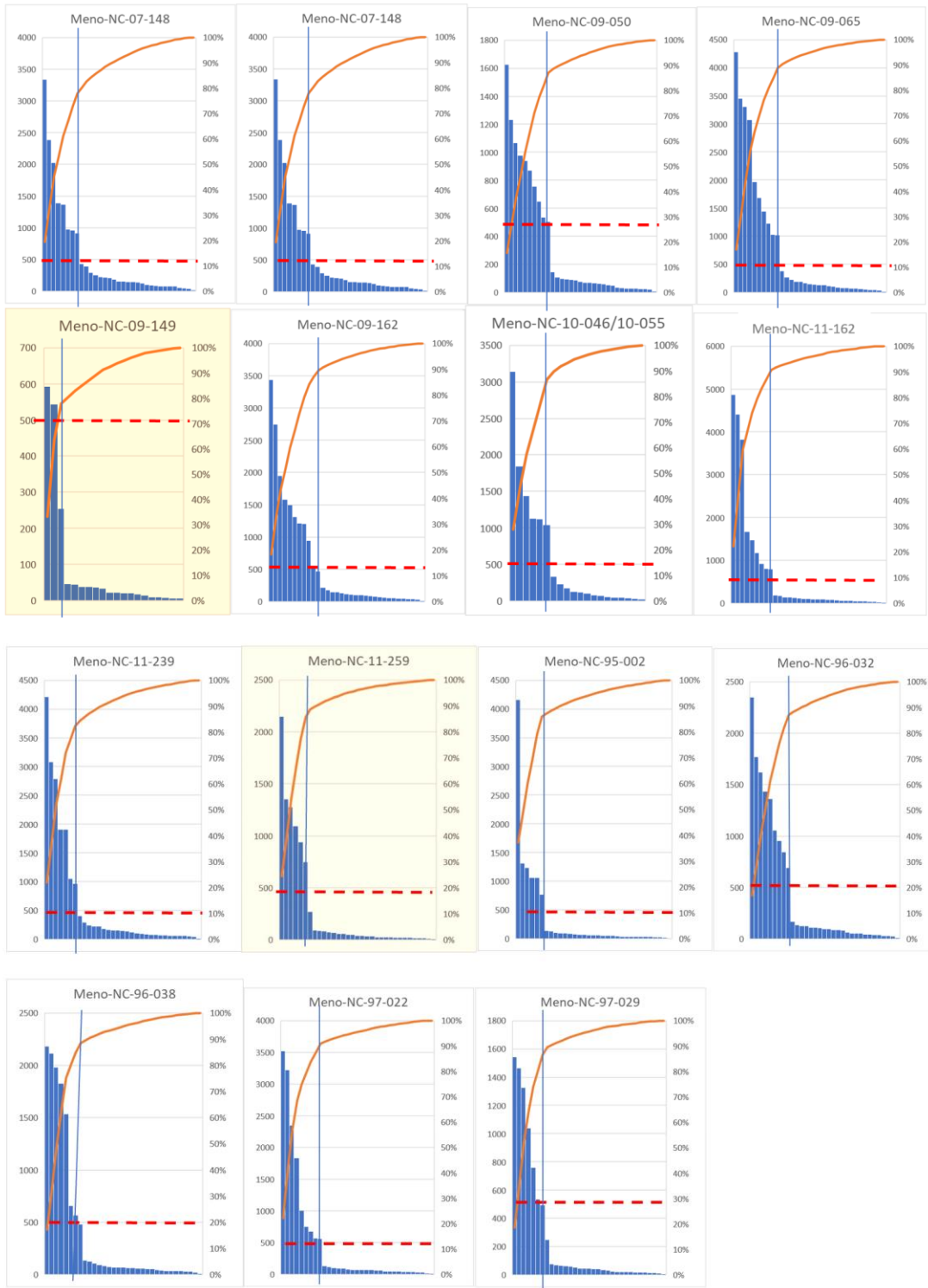


MOLECULAR ECOLOGY RESOURCES

Humpback whales

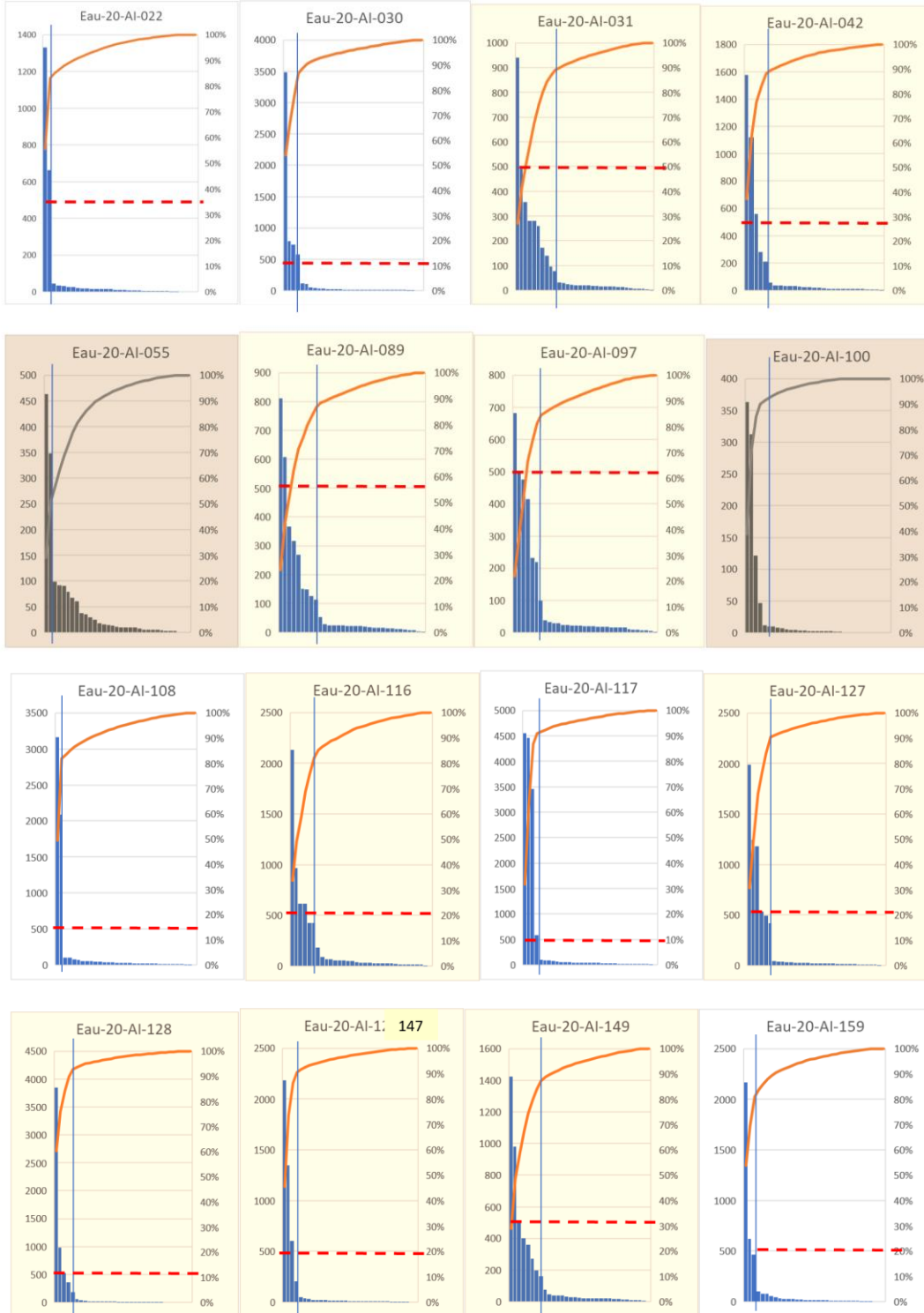


MOLECULAR ECOLOGY RESOURCES

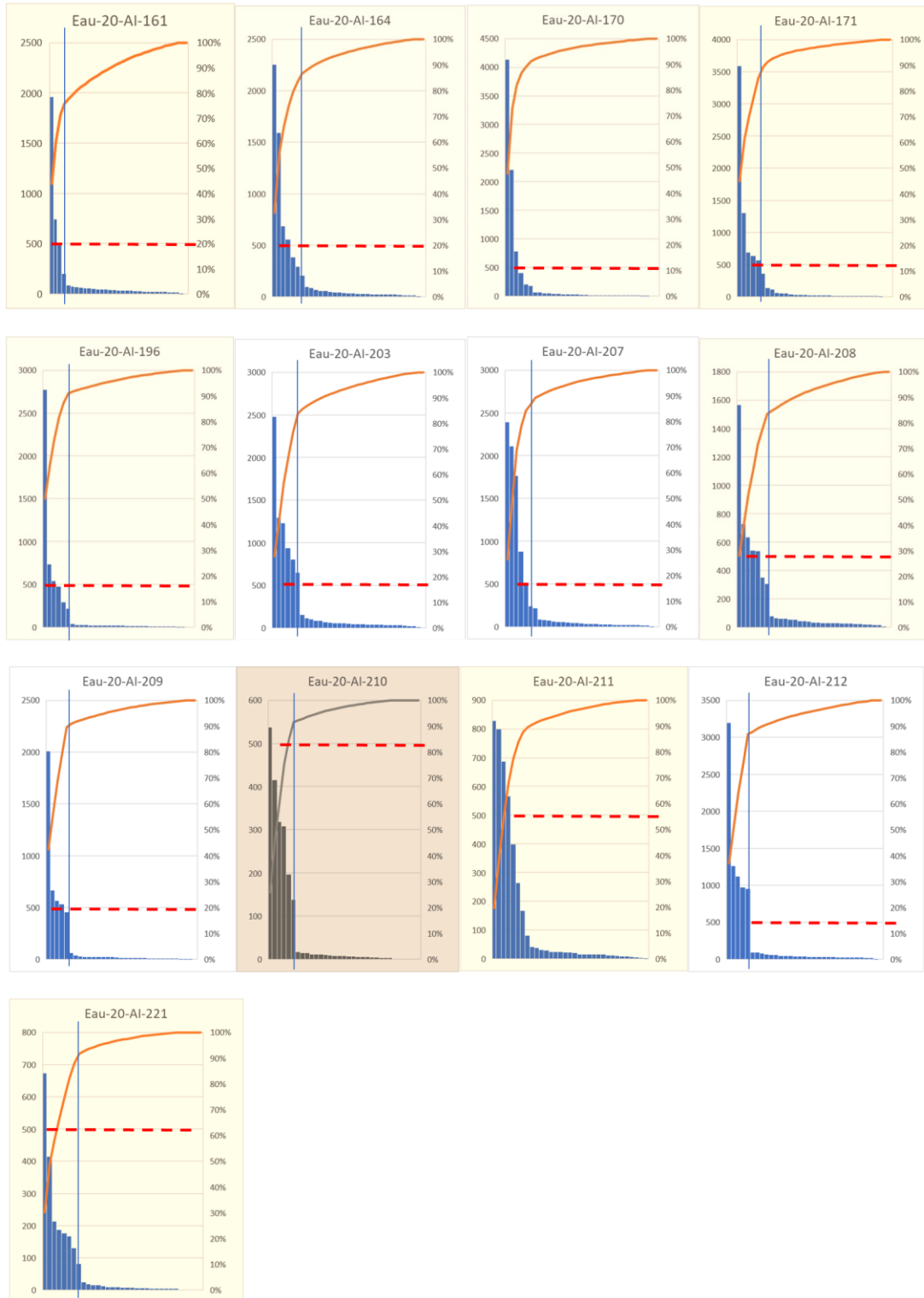


MOLECULAR ECOLOGY RESOURCES

Southern right whales



MOLECULAR ECOLOGY RESOURCES

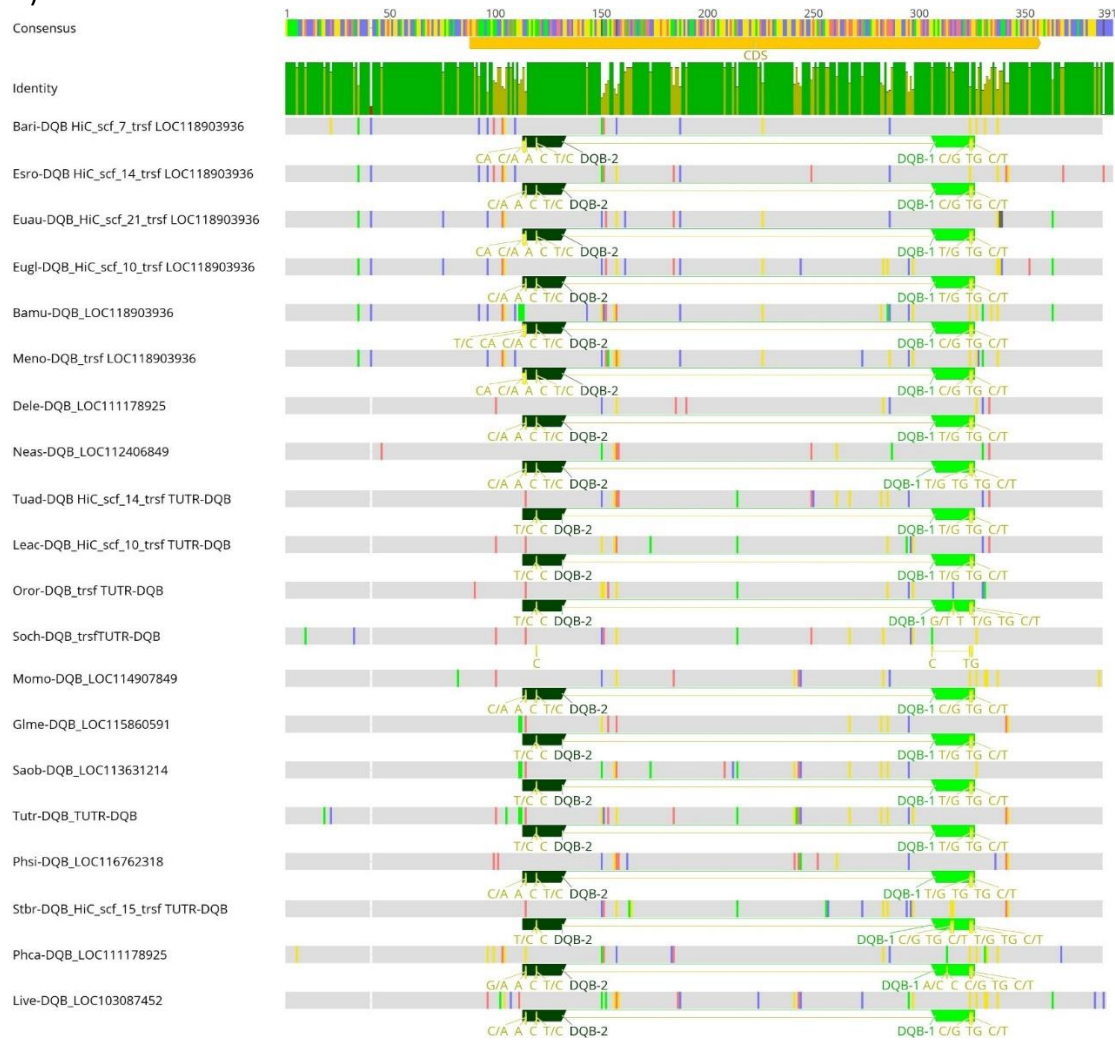


MOLECULAR ECOLOGY RESOURCES

Figure S4: DQB and DRB alignment of genes at amplicon length from genomes

A) DQB and B) DRB alignment of genes at amplicon length from genomes (see accession numbers Table 1S) with exon 2 coding sequence indicated in orange in the consensus sequence. Primer binding sites for DQB2 (forward; dark green) and DQB1 (reverse, light green) primers for each sequence are shown, respectively, under the sequence. These are only shown if the primers bind to the sequence with a maximum of four mismatches. Mismatches of the primer binding sequence are shown in yellow within the primer. Nucleotide bases are coloured (A = red, C = blue, G = yellow, T = green). Grey indicates the same nucleotide base as the consensus and only mismatches to the consensus sequence are shown. Mean pairwise identity over all sequences is shown in green where (Green: 100% identity; Greeny-brown: at least 30% and under 100% identity; Red: below 30% identity). DRB primer binding sites are also shown in B) for DRB-a gene sequences.

A)



MOLECULAR ECOLOGY RESOURCES

B)

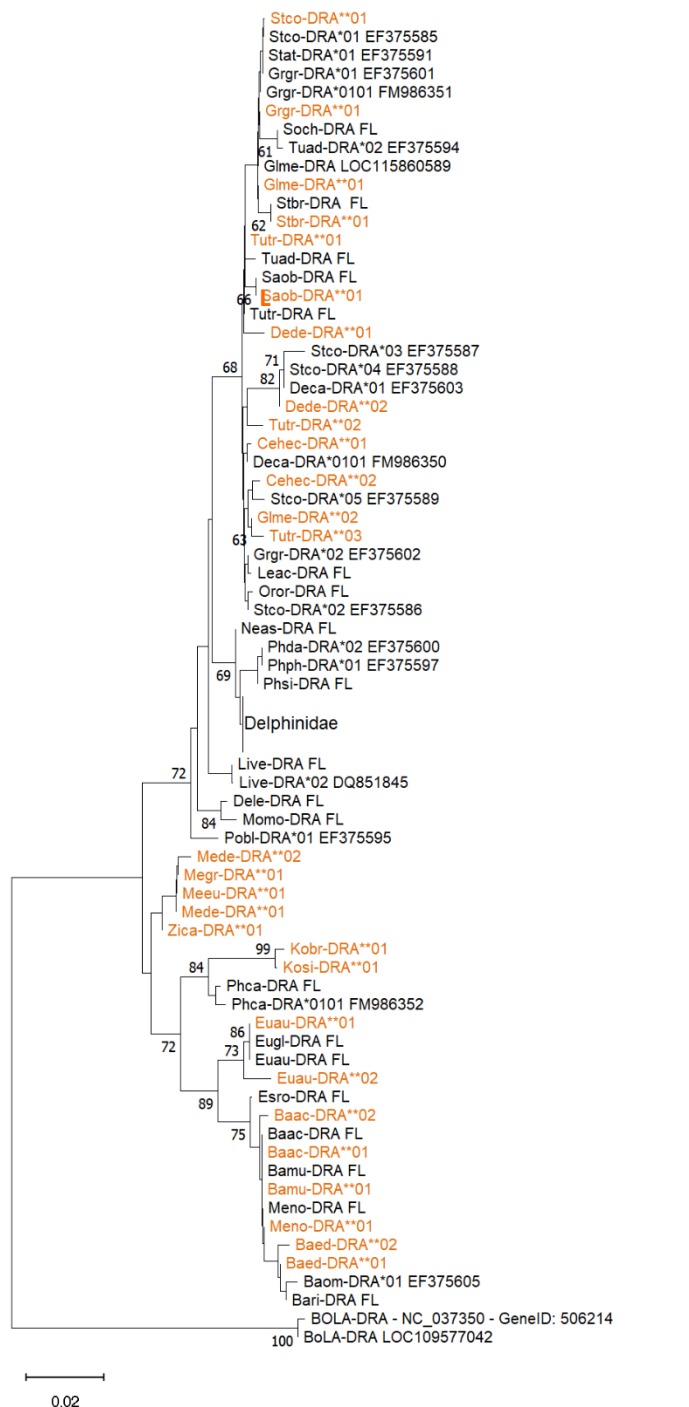


MOLECULAR ECOLOGY RESOURCES

Figure S5: Neighbor-Joining phylogenetic trees for amplicon length genes for A) DRA, B) DRB, C) DQA, and D) DQB.

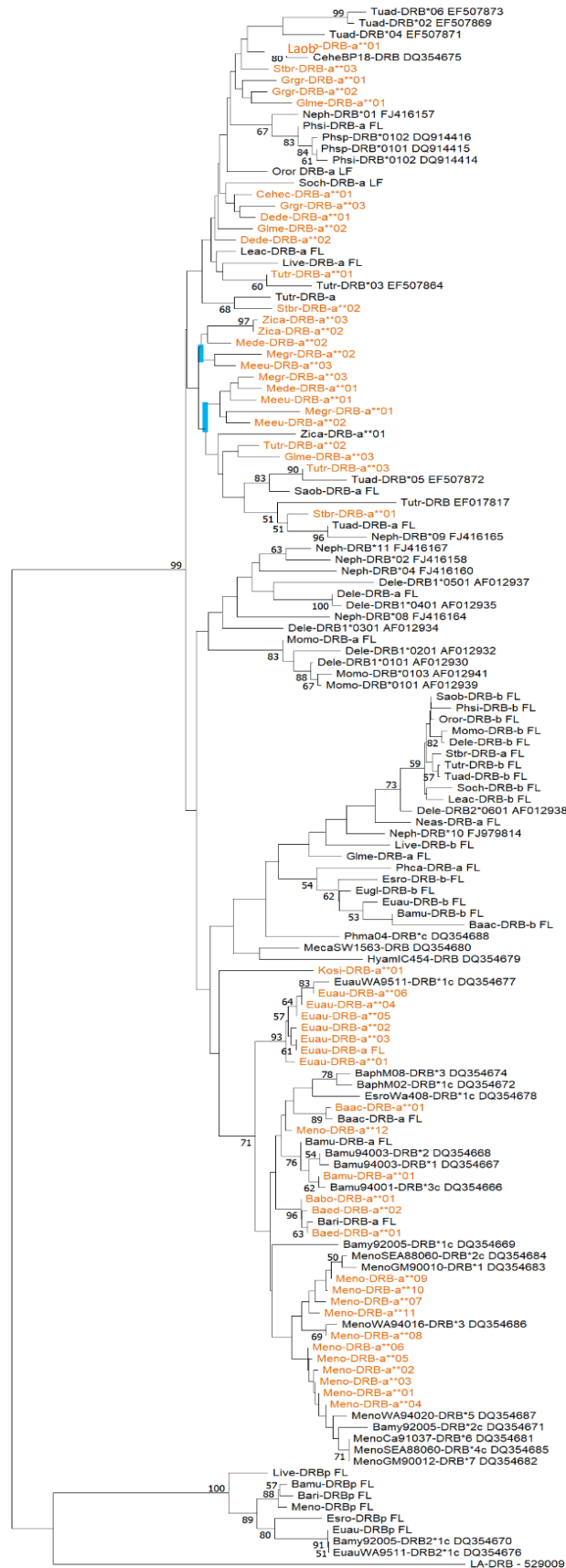
Alleles from genome assemblies are labelled FL, alleles found in our study are marked with ** and in red while alleles found in Genbank contain their Genbank accession number. Accession numbers for genome assemblies can be found in Table S1. Light blue indicates toothed whales (odontocetes) and dark blue for baleen whales (mysticetes).

A)



MOLECULAR ECOLOGY RESOURCES

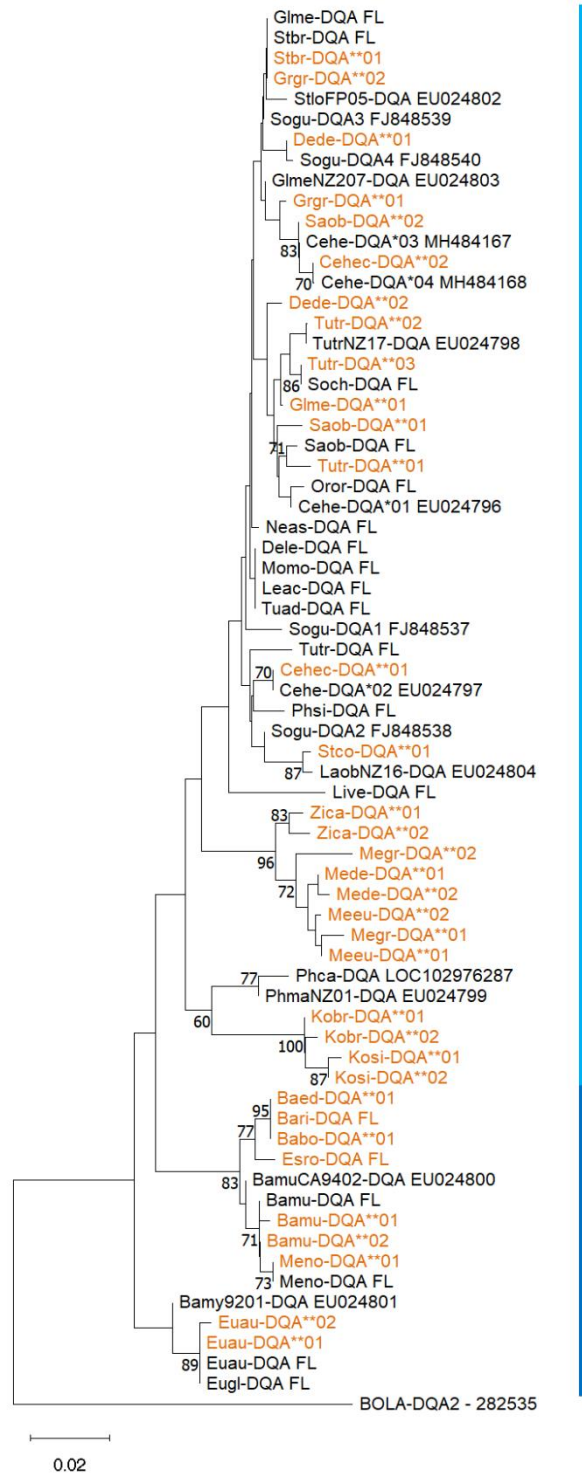
B)



0.050

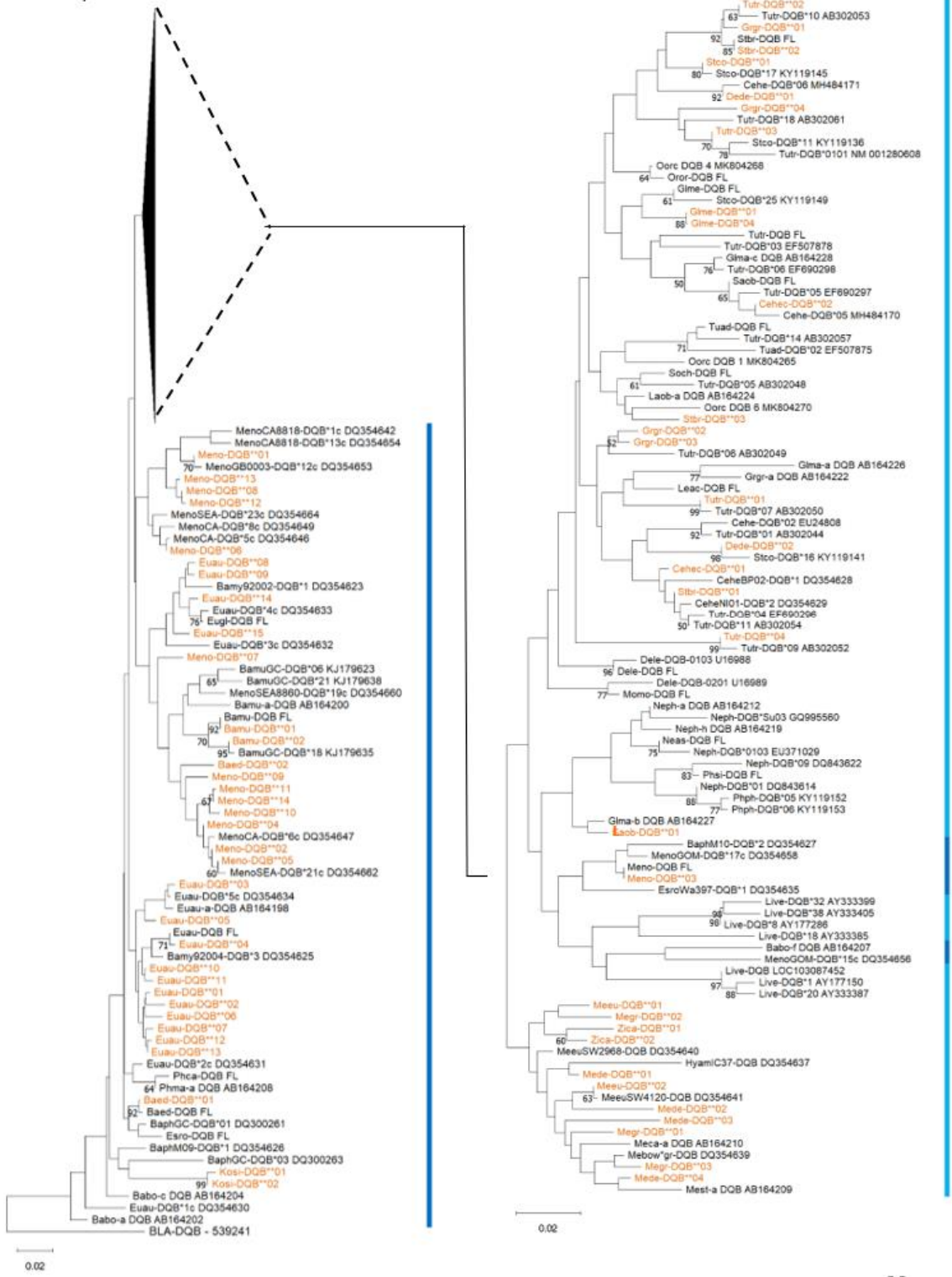
MOLECULAR ECOLOGY RESOURCES

C)



MOLECULAR ECOLOGY RESOURCES

D)



MOLECULAR ECOLOGY RESOURCES

Acknowledgements and references for genome assemblies as provided on DNA-Zoo

Genome assemblies from 25 species were downloaded from DNA Zoo and investigated for MHC class I and class IIa regions. Five of those assemblies (shaded in red) were not used further due to fragmented or missing MHC regions. DNA Zoo genomes were scaffolded with Hi-C data using 3D-DNA (Dudchenko et al., 2017) and reviewed in Juicebox Assembly Tools (Dudchenko et al., 2018). More details on the methodology are available at www.dnazoo.org/methods.

Common name	Organism name	four letter code	Database	Reference
North Atlantic right whale	<i>Eubalaena glacialis</i>	Eugl	DNA Zoo	https://www.dnazoo.org/assemblies/Eubalaena_glacialis 1)
Southern right whale	<i>Eubalaena australis</i>	Euau	DNA Zoo	https://www.dnazoo.org/assemblies/Eubalaena_australis 2)
Fin whale	<i>Balaenoptera physalus</i>	Baph	DNA Zoo	https://www.dnazoo.org/assemblies/Balaenoptera_physalus 3)
Gray whale	<i>Eschrichtius robustus</i>	Esro	DNA Zoo	https://www.dnazoo.org/assemblies/Eschrichtius_robustus 4)
Humpback whale	<i>Megaptera novaeangliae</i>	Meno	DNA Zoo	https://www.dnazoo.org/assemblies/Megaptera_novaeangliae 5)
Rice's whale	<i>Balaenoptera ricei</i>	Bari	DNA Zoo	https://www.dnazoo.org/assemblies/Balaenoptera_ricei 6)
Atlantic spotted dolphin	<i>Stenella frontalis</i>	Stfr	DNA Zoo	https://www.dnazoo.org/assemblies/Stenella_frontalis 7)
Atlantic white-sided dolphin	<i>Lagenorhynchus acutus</i>	Laac	DNA Zoo	https://www.dnazoo.org/assemblies/Lagenorhynchus_acutus 8)
Clymene dolphin	<i>Stenella clymene</i>	Stcl	DNA Zoo	https://www.dnazoo.org/assemblies/Stenella_clymene 9)
Commerson's dolphin	<i>Cephalorhynchus commersonii</i>	Ceco	DNA Zoo	https://www.dnazoo.org/assemblies/Cephalorhynchus_commersonii 10)
Bottlenose dolphin	<i>Tursiops truncatus</i>	Tutr	DNA Zoo	https://www.dnazoo.org/assemblies/Tursiops_truncatus 11)
Eastern spinner dolphin	<i>Stenella longirostris</i>	Stlo	DNA Zoo	https://www.dnazoo.org/assemblies/Stenella_longirostris_orientalis 12)
Indo-Pacific bottlenose dolphin	<i>Tursiops aduncus</i>	Tuad	DNA Zoo	https://www.dnazoo.org/assemblies/Tursiops_aduncus 13)
Killer whale	<i>Orcinus orca</i>	Oror	DNA Zoo	https://www.dnazoo.org/assemblies/Orcinus_orca 14)
Long-finned pilot whale	<i>Globicephala melas</i>	Glme	DNA Zoo	https://www.dnazoo.org/assemblies/Globicephala_melas 15)
Pacific white-sided dolphin	<i>Sagmatias obliquidens</i>	Saab	DNA Zoo	https://www.dnazoo.org/assemblies/Lagenorhynchus_obliquidens 16)
Pantropical spotted dolphin	<i>Stenella attenuata</i>	Stat	DNA Zoo	https://www.dnazoo.org/assemblies/Stenella_attenuata 17)
Risso's dolphin	<i>Grampus griseus</i>	Grgr	DNA Zoo	https://www.dnazoo.org/assemblies/Grampus_griseus 18)
Rough-toothed dolphin	<i>Steno bredanensis</i>	Stbr	DNA Zoo	https://www.dnazoo.org/assemblies/Steno_bredanensis 19)
Pygmy sperm whale	<i>Kogia breviceps</i>	Kobr	DNA Zoo	https://www.dnazoo.org/assemblies/Kogia_breviceps 20)
Beluga	<i>Delphinapterus leucas</i>	Dele	DNA Zoo	https://www.dnazoo.org/assemblies/Delphinapterus_leucas 21)
Harbor porpoise	<i>Phocoena phocoena</i>	Phoh	DNA Zoo	https://www.dnazoo.org/assemblies/Phocoena_phocoena 22)
Blainville's beaked whale	<i>Mesoplodon densirostris</i>	Mede	DNA Zoo	https://www.dnazoo.org/assemblies/Mesoplodon_densirostris 23)
Gervais' beaked whale	<i>Mesoplodon europaeus</i>	Meeu	DNA Zoo	https://www.dnazoo.org/assemblies/Mesoplodon_europaeus 24)
Stejneger's beaked whale	<i>Mesoplodon stejnegeri</i>	Mest	DNA Zoo	https://www.dnazoo.org/assemblies/Mesoplodon_stejnegeri 25)

Mysticetes

1) North Atlantic right whale

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al., 2017), see (Dudchenko et al., 2018) for details. This work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen used in this study was collected by NOAA/T. Rowles/B. Bonde and provided by the National Marine Mammal Tissue Bank, which is maintained by the National Institute of Standards and Technology (NIST) in the NIST Biorepository, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program.

2) Southern Right whale

This assembly was generated in collaboration with the University of Auckland.

MOLECULAR ECOLOGY

RESOURCES

3) *Fin whale*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. The sample for the assembly was collected by The Marine Mammal Center under the Marine Mammal Health and Stranding Program (MMHSPR) Permit No. 18786-04 issued by the National Marine Fisheries Service (NMFS) in accordance with the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The work at DNA Zoo was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03.

4) *Gray whale*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. Work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen used in this study was collected by Pam Tuomi (Alaska Sealife Center) from Girdwood, Alaska and provided by the National Marine Mammal Tissue Bank, which is maintained by the National Institute of Standards and Technology (NIST) in the NIST Biorepository, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program and the Alaska Marine Mammal Tissue Archival Project.

5) *Humpback whale*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details.

6) *Rice's whale*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen(s) used in this study were collected by the NMFS Panama City Lab/Lydia Staggs and provided by the National Marine Mammal Tissue Bank, which is maintained in the Marine Environmental Specimen Bank (Marine ESB) at NIST, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program.

Odontocetes

7) *Atlantic spotted dolphin*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This work was performed under Marine Mammal Health and Stranding Response Program

MOLECULAR ECOLOGY RESOURCES

(MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen used in this study was collected by Wayne McFee and Jessica Conwa (NOAA) and provided by the National Marine Mammal Tissue Bank, which is maintained by the National Institute of Standards and Technology (NIST) in the NIST Biorepository, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program.

8) *Atlantic white-sided dolphin*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen used in this study was collected by Wayne McFee (NOAA) and provided by the National Marine Mammal Tissue Bank, which is maintained by the National Institute of Standards and Technology (NIST) in the NIST Biorepository, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program.

9) *Clymene dolphin*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen used in this study (Field ID WAM 602; Storage ID NM15K707C) was collected from from Topsail Island, NC by Bill McLellan (UNCW), and provided by the National Marine Mammal Tissue Bank, which is maintained in the Marine Environmental Specimen Bank (Marine ESB) at NIST, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program.

10) *Commerson's dolphin*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details.

11) *Bottlenose dolphin*

Neely, B.A. "The genome of the Atlantic bottlenose dolphin (*Tursiops truncatus*)."
Unpublished. <https://www.ncbi.nlm.nih.gov/nuccore/MRVK00000000.1>

12) *Eastern Spinner dolphin*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details.

MOLECULAR ECOLOGY RESOURCES

This work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen used in this study was collected by Wayne McFee and Jessica Conwa (NOAA) and provided by the National Marine Mammal Tissue Bank, which is maintained by the National Institute of Standards and Technology (NIST) in the NIST Biorepository, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program.

13) Indo-Pacific bottlenose dolphin

Vijay, Nagarjun, Chungoo Park, Jooseong Oh, Soyeong Jin, Elizabeth Kern, Hyun Woo Kim, Jianzhi Zhang, and Joong-Ki Park. 2018. "Population Genomic Analysis Reveals Contrasting Demographic Changes of Two Closely Related Dolphin Species in the Last Glacial." *Molecular Biology and Evolution* 35 (8): 2026-33. <https://doi.org/10.1093/molbev/msy108>.

14) Killer whale

Foote, Andrew D., Yue Liu, Gregg W. C. Thomas, Tomas Vinar, Jessica Alfoldi, Jixin Deng, Shannon Dugan, et al. 2015. "Convergent Evolution of the Genomes of Marine Mammals." *Nature Genetics* 47 (3): 272-75. <https://doi.org/10.1038/ng.3198>.

15) Long-finned pilot whale

Culibrk,L., Leelakumari,S., Taylor,G.A., Tse,K., Cheng,D., Chuah,E., Kirk,H., Pandoh,P., Troussard,A., Zhao,Y., Mungall,A., Moore,R., Bourque,L., Daoust,P.-Y., Jones,M., Ortenburger,A., Marra,M.A. and Jones,S.J.M. "The genome of the Pilot Whale (*Globicephala melas*)." Unpublished. <https://www.ncbi.nlm.nih.gov/nuccore/SWEB00000000.1>.

16) Pacific white-sided dolphin

Culibrk,L., Leelakumari,S., Taylor,G.A., Tan,R., Tse,K., Cheng,D., Chuah,E., Kirk,H., Macleod,T., Pandoh,P., Troussard,A., Zhao,Y., Mungall,A.J., Moore,R., Marra,M.A., Haulena,M. and Jones,S.J.M. "The genome of the whitesided dolphin." Unpublished. <https://www.ncbi.nlm.nih.gov/nuccore/RCWK00000000.1>.

17) Pantropical spotted dolphin

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen(s) used in this study were collected by the NMFS Panama City Lab/Lydia Staggs and provided by the National Marine Mammal Tissue Bank, which is maintained in the Marine Environmental Specimen Bank (Marine ESB) at NIST, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program.

MOLECULAR ECOLOGY RESOURCES

18) *Risso's dolphin*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This sample was collected by The Marine Mammal Center under the Marine Mammal Health and Stranding Program (MMHSPR) Permit No. 18786-04 issued by the National Marine Fisheries Service (NMFS) in accordance with the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The work at DNA Zoo was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03.

19) *Rough-toothed dolphin*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This sample was collected by The Marine Mammal Center under the Marine Mammal Health and Stranding Program (MMHSPR) Permit No. 18786-04 issued by the National Marine Fisheries Service (NMFS) in accordance with the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The work at DNA Zoo was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03.

20) *Pygmy sperm whale*

Genevoux, D.P., Serres, A., Armstrong, J., Johnson, J., Marinescu, V.D., Murén, E., Juan, D., Bejerano, G., Casewell, N.R., Chemnick, L.G., Damas, J., Di Palma, F., Diekhans, M., Fiddes, I.T., Garber, M., Gladyshev, V.N., Goodman, L., Haerty, W., Houck, M.L., Hubley, R., Kivioja, T., Koepfli, K.-P., Kuderna, L.F.K., Lander, E.S., Meadows, J.R.S., Murphy, W.J., Nash, W., Noh, H.J., Nweeia, M., Pfenning, A.R., Pollard, K.S., Ray, D.A., Shapiro, B., Smit, A.F.A., Springer, M.S., Steiner, C.C., Swofford, R., Taipale, J., Teeling, E.C., Turner-Maier, J., Alföldi, J., Birren, B., Ryder, O.A., Lewin, H.A., Paten, B., Marques-Bonet, T., Lindblad-Toh, K., Karlsson, E.K., Zoonomia Consortium, 2020. A comparative genomics multitool for scientific discovery and conservation. *Nature* 587, 240–245. <https://doi.org/10.1038/s41586-020-2876-6>.

21) *Beluga*

Jones, Steven J. M., Gregory A. Taylor, Simon Chan, Rene L. Warren, S. Austin Hammond, Steven Bilobram, Gideon Mordecai, et al. 2017. "The Genome of the Beluga Whale (*Delphinapterus leucas*)." *Genes* 8 (12): 378. <https://doi.org/10.3390/genes8120378>.

22) *Harbor porpoise*

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen used in this study was collected by Carol A. Stephens and provided by the National Marine Mammal Tissue Bank, which is maintained by the National Institute of Standards and Technology (NIST) in the NIST Biorepository, and which is operated under the

MOLECULAR ECOLOGY

RESOURCES

direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program and the Alaska Marine Mammal Tissue Archival Project.

23) Blainville's beaked whale

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen used in this study was collected by Wayne McFee (NOAA) and provided by the National Marine Mammal Tissue Bank, which is maintained by the National Institute of Standards and Technology (NIST) in the NIST Biorepository, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program.

24) Gervais' beaked whale

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen used in this study was collected by NOAA Wayne McFee and provided by the National Marine Mammal Tissue Bank, which is maintained by the National Institute of Standards and Technology (NIST) in the NIST Biorepository, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program.

25) Stejneger's beaked whale

The draft assembly was generated by the DNA Zoo team from short insert-size PCR-free DNA-Seq data using w2rap-contigger (Clavijo et al. 2017), see (Dudchenko et al., 2018) for details. This work was performed under Marine Mammal Health and Stranding Response Program (MMHSRP) Permit No. 18786-03 issued by the National Marine Fisheries Service (NMFS) under the authority of the Marine Mammal Protection Act (MMPA) and Endangered Species Act (ESA). The specimen used in this study was collected by Pam Tuomi (ASLC) and provided by the National Marine Mammal Tissue Bank, which is maintained by the National Institute of Standards and Technology (NIST) in the NIST Biorepository, and which is operated under the direction of NMFS with the collaboration of USGS, USFWS, MMS, and NIST through the Marine Mammal Health and Stranding Response Program and the Alaska Marine Mammal Tissue Archival Project.

MOLECULAR ECOLOGY RESOURCES

References

- Clavijo, B. J., Accinelli, G. G., Wright, J., Heavens, D., Barr, K., Yanes, L., & Di-Palma, F. (2017). W2RAP: a pipeline for high quality, robust assemblies of large complex genomes from short read data. *BioRxiv*, 110999.
<https://www.biorxiv.org/content/10.1101/110999v1><https://www.biorxiv.org/content/10.1101/110999v1.abstract>
- Dudchenko, O., Batra, S. S., Omer, A. D., Nyquist, S. K., Hoeger, M., Durand, N. C., Shamim, M. S., Machol, I., Lander, E. S., Aiden, A. P., & Aiden, E. L. (2017). De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science*, 356(6333), 92–95.
<https://doi.org/10.1126/science.aal3327>
- Dudchenko, O., Shamim, M. S., Batra, S. S., Durand, N. C., Musial, N. T., Mostofa, R., Pham, M., Hilaire, B. G. S., Yao, W., Stamenova, E., Hoeger, M., Nyquist, S. K., Korchina, V., Pletch, K., Flanagan, J. P., Tomaszewicz, A., McAloose, D., Estrada, C. P., Novak, B. J., ... Nathaniel, T. (2018). The Juicebox Assembly Tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under \$1000. *BioRxiv*, 254797.
<https://www.biorxiv.org/content/early/2018/01/28/254797>https://www.biorxiv.org/content/early/2018/01/28/254797?utm_content=buffer421d5&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer