

Technical Report (not peer reviewed)

A method to investigate close kin in wildlife populations: the case of the southern right whale (*Eubalaena australis*) in the Indian sector of the Antarctic Ocean

Megumi TAKAHASHI* and Luis A. PASTENE

Institute of Cetacean Research, 4–5 Toyomi-cho, Chuo-ku, Tokyo 104–0055, Japan

*Contact e-mail: takahashi@cetacean.jp

ABSTRACT

This paper describes a method for investigating close kin in wildlife species, which is based on LOD (Logarithm of the odds) score—a relatedness index for a pair of genotypes. The method was applied to study the distribution of close kin in southern right whales in the feeding grounds of the Indo-Pacific sector of the Antarctic Ocean, using biopsy samples of 183 individual whales collected during Japanese and International Whaling Commission’s whale surveys in the Antarctic. The description of the method and results for southern right whales in this paper are based on a recently published paper derived from a research collaboration between scientists of the Institute of Cetacean Research and the University of Bergen, Norway.

INTRODUCTION

The genetic determination of close kin pairs has been proposed as a general framework for estimating key population parameters, such as population size and growth rate in wild animals (Bravington *et al.*, 2016). Also, kinship information can assist the interpretation of movement and social and population structure in whales (Skaug *et al.*, 2010). Therefore, information on close kin has the potential to contribute to the conservation of whale species through the monitoring of important population parameters. See Taguchi (2020) for a review of genetic methods to infer kinship inferences in whales.

A commonly used test statistic for close kin inference is the LOD (Logarithm of the odds) score, which is readily calculated from a pair of DNA profiles (Meagher, 1986; Skaug, 2001). Two individuals may be classified as related, with no further specification of the type of relationship if their LOD score exceeds a predefined critical value.

The objective of this paper is to present a brief outline of the analytical methodology of close kinship inference based on LOD. This paper also presents the results of the application of this methodology on southern right whales in the feeding grounds of the Indo-Pacific sector of the Antarctic. The description of the method and results for southern right whales in this paper are based on a recently published

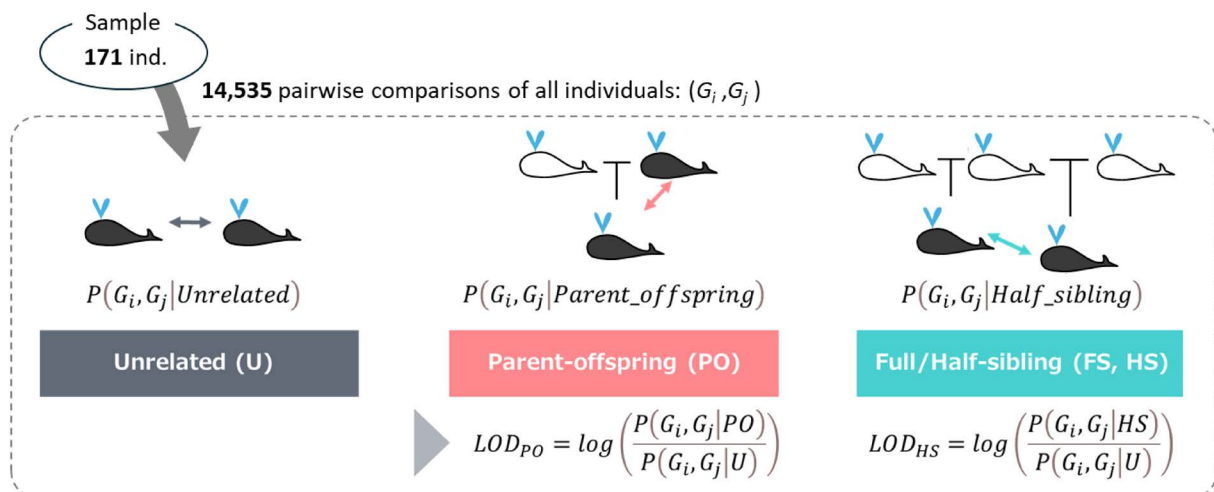


Figure 1. Examples of relatedness index for a pair of genotype for a case of 171 individuals analyzed.

paper derived from a research collaboration between scientists of the Institute of Cetacean Research (ICR) and the University of Bergen, Norway (Takahashi *et al.*, 2024).

ANALYTICAL PROCEDURES

LOD score analysis

The LOD is defined as the log-likelihood ratio of the probability of an observed genotype pair when they share a certain relationship, e.g., parent-offspring, relative to the probability of the genotype pair when they are unrelated. For the parent-offspring relationship (PO), this amounts to

$$LOD_{PO}(i, j) = \ln \left(\frac{P(G_i, G_j | \text{Parent} - \text{offspring})}{P(G_i, G_j | \text{Unrelated})} \right).$$

Here, G_i and G_j are the observed genotypes for two specimens, i and j . LOD scores for other kinship categories are defined similarly (Figure 1).

The observed genotypes may contain genotyping errors. To address this, a simple error model with a constant and independent per-allele error rate can be used. For a parent-offspring pair and self/monozygotic relationships, an error model ensures that the LOD score is well defined for all genotype pairs. Otherwise, the LOD score would be undefined for all pairs containing loci with no compatible alleles.

The second-degree relationships of half-siblings (HS),

grandparents, and full aunts and uncles all have the same LOD score and cannot be distinguished on the basis of genotypes. We refer to this as the LOD_{HS} (half-sibling LOD score). It should be kept in mind that LOD_{HS} does not necessarily refer to only half-siblings. We define the maximum likelihood kinship for a genotype pair as the kinship category that gives the highest probability for the observed genotype, i.e., the relationship with the highest $P(G_i, G_j | \text{Kinship})$. This is equivalent to selecting the kinship category with the highest LOD score, since its denominator is the same for all kinship categories.

APPLICATION OF THE METHOD TO THE SOUTHERN RIGHT WHALE

Background of the southern right whale

Southern right whales (*Eubalaena australis*) are widely distributed across the three ocean basins in the Southern Hemisphere: South Atlantic, Indian Ocean, and South Pacific, mainly between latitudes 16°S and 65°S. Southern right whales approach the continental coasts and some islands for breeding, calving and resting during the austral winter and early spring. The primary breeding grounds of this species are located in the waters off South Africa, South West Australia, mainland New Zealand, New Zealand Sub-Antarctic and Argentina (Figure 2) (Carroll *et al.*, 2014; 2016; Cranswick, 2022; IWC, 2001; 2013). Also, southern right whales occur in winter in coastal areas of

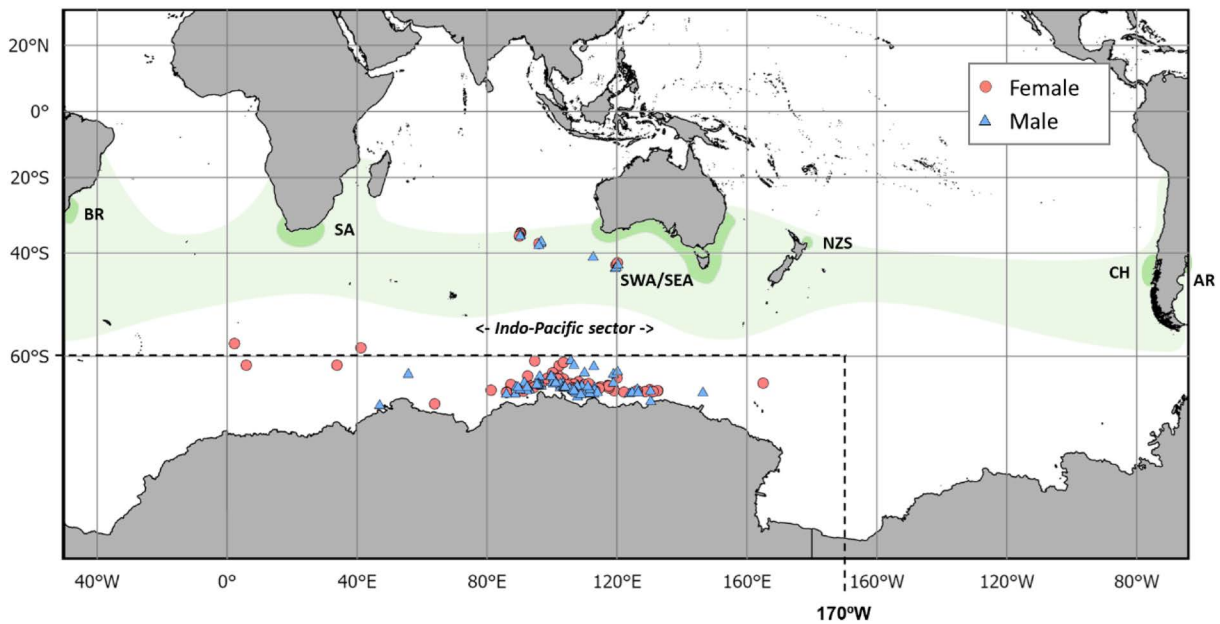


Figure 2. Historical distribution and primary calving grounds of southern right whales. The geographical positions of the sampled whales for this study are shown by sex. Surveys were conducted in the austral summer seasons (December to March) from 1993/94 to 2018/19. The map was obtained in the R statistical environment (R core Team, 2023) using 'rnaturalearth' package (Massicotte and South, 2023), and the world vector map data was obtained from Natural Earth (public domain): <http://www.naturalearthdata.com>.

South Eastern Australia, Chile and Brazil.

Previous studies on distribution and movement based on photo-id matches (Bannister, 1999), historical whaling data and sighting surveys (Bannister, 2001) showed that whales in the Indian sector of the Antarctic (85°–135°E) are associated with breeding grounds in the Australasian regions. Also, a previous study on mtDNA analysis supported a close relationship between whales in the Antarctic Indian sector and whales in the Southwestern Australia breeding ground (Pastene *et al.*, 2018).

Biopsy samples

A total of 183 skin/blubber biopsy samples were obtained opportunistically from free-ranging southern right whales along the sighting surveys of the Japanese Whale Research Program under Special Permit in the Antarctic (JARPA/JARPAII), the International Whaling Commission-International Decade for Cetacean Research/Southern Ocean Whale and Ecosystem Research (IWC-IDCR/SOWER) programs, and the New Scientific Whale Research Program in the Antarctic Ocean (NEWREP-A) in the Antarctic Ocean from 60°W to 170°W across the 180° meridian, south of 35°S during the austral summer seasons (December to March) 1993/94 to 2018/19.

Most of the samples were taken in the Antarctic Indian sector (85°–135°E), south of 60°S from 1993/94 to 2015/16. Figure 2 shows the geographical distribution of the southern right whales sampled. For each sample, sampling date and geographical location were available. In some cases, visually estimated measurements of body lengths of the animals sampled were recorded in conjunction with other ancillary information such as the presence of calf. Visual estimates of body length were made from the vessels by experienced scientists and crew members. This additional non-genetic information

assisted the interpretation of the results of close kinship inferences.

DNA profile preparation

Microsatellite DNA (msDNA) genotypes were the primary data used for estimating close kinship. Also, the sex information and mitochondrial DNA (mtDNA) haplotypes were used only to assist the interpretation of close kinship inferences. The DNA samples were genotyped using 14 microsatellite loci: EV1Pm, EV14Pm, EV21Pm, EV37Mn, EV94Mn (Valsecchi and Amos, 1996), GT023, GT211, GT310 (Bérubé *et al.*, 2009), GATA028 (Palsbøll *et al.*, 1997), DlrFCB17 (Buchanan *et al.*, 1996), TR3G2, TR2G5, TR2F2, and TR3F3 (Frasier *et al.*, 2006). The SRY locus located on the Y chromosome was also used for sex determination following the method of Abe *et al.* (2001) with a slight modification described in Pastene *et al.* (2022).

Also, the first 470 base pairs (bp) at the 5' end of the mtDNA control region was sequenced. The mtDNA haplotypes were used for interpretation purposes.

Results of the close kinship inference

Determination of the cut-off point value

Figure 3 shows the LOD score distributions (LOD_{p0}) for simulated and observed data sets. There was an overlap between the three distributions simulated from different kinship categories (Figure 3, left). Table 1 shows the number of detected close kin dyads at various levels of the cut-off point of LOD score, as well as the expected number of false positives from unrelated pairs. The challenge is to find a way to determine the value of the cut-off point that will maximize the number of correctly inferred pairs of related individuals, while keeping the number of pairs incorrectly inferred as close relatives to a reason-

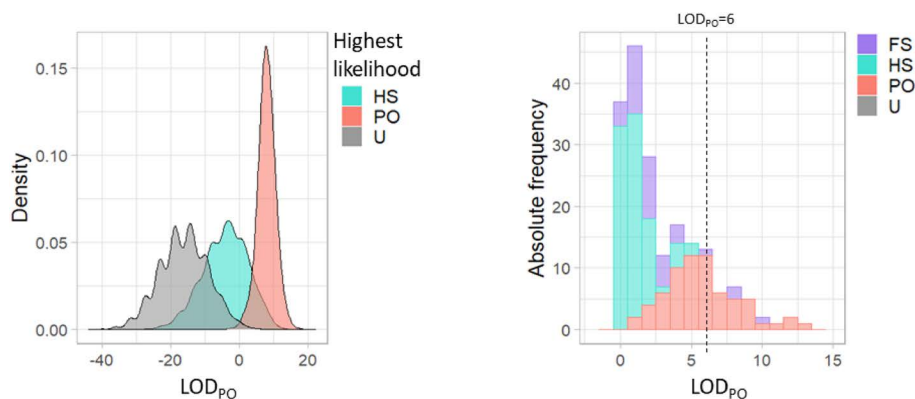


Figure 3. Simulated (left) and observed (right) distribution of the LOD_{p0} scores. For the three simulated datasets (HS, PO, U) normalized densities (unit area) are shown, while for the real data the absolute frequency of each relationship category, as assigned by the maximum likelihood, is displayed.

Table 1

Number of related pairs by different (maximum probability) kinship category (PO/HS/FS) for LOD_{PO} score, for different cut-off LOD values.

Kinship categories for LOD score	Cut-off point	# highest likelihood (# expected FP)					# expected FN on category "PO"
		FS	HS	PO			
HS	8	1	—	—	—	2	—
	7	1	—	—	—	4	(0.01)
	6	2	(0.09)	—	(0.16)	8	(0.19)
	5	5	(1.00)	6	(1.00)	15	(0.97)
	4	9	(4.00)	14	(8.00)	25	(4.00)
	3	28	(18.00)	41	(49.00)	43	(12.00)
PO	8	2	(0.07)	—	—	13	(0.60)
	7	3	(0.16)	—	—	18	(1.00)
	6	3	(0.29)	—	(0.01)	25	(4.00)
	5	4	(0.74)	1	(0.31)	39	(9.00)
	4	6	(1.00)	4	(1.00)	51	(16.00)
	3	9	(3.00)	7	(4.00)	58	(23.00)

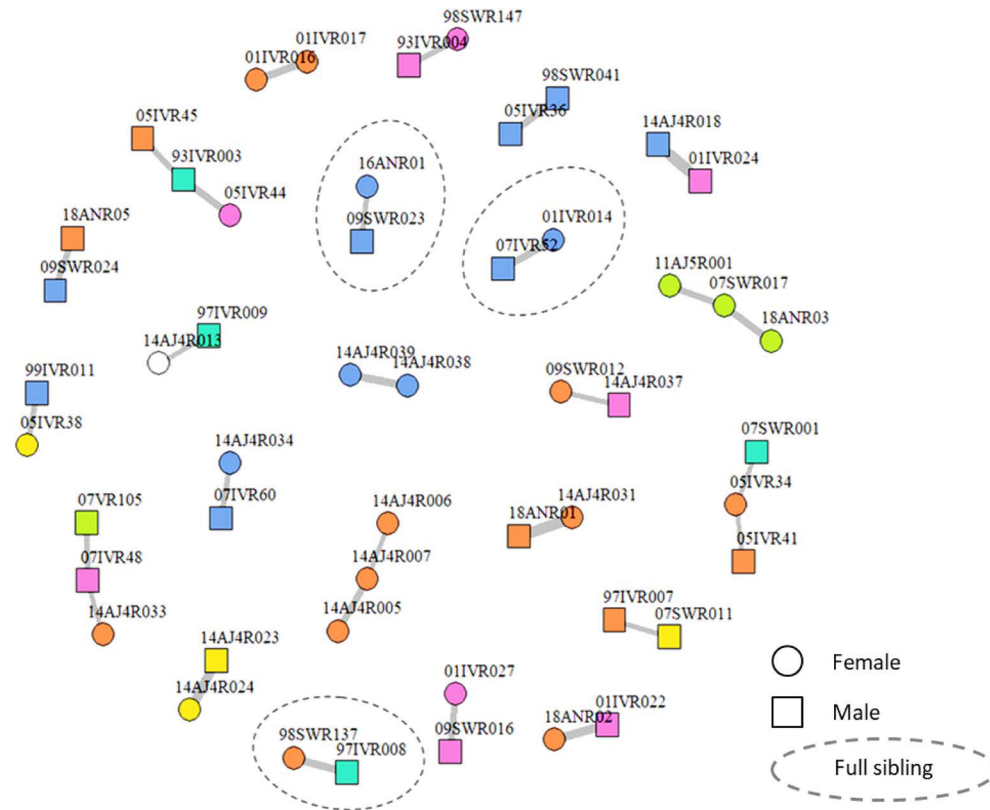


Figure 4. Schematic representation of the dyads and triads in Table 1 ($LOD_{PO} > 6$). The strength of evidence (LOD_{PO}) is indicated by the width of the line connecting related individuals. Mitochondrial DNA haplotypes are shown by color: circles denote females and squares denote males. The dashed circles show dyads inferred as full-sibling pairs by maximum likelihood.

able minimum. In this case, the best cut-off point under this criterion was $LOD_{PO} > 6$.

Identification of close kin dyads

Figure 4 provides a schematic representation of the information of the 28 dyads with $LOD_{PO} > 6$. They consist of a

total of 25 possible PO pairs and three possible full-sibling (FS) pairs, as determined by maximum likelihood classification. Some additional genetic (mtDNA haplotype) and biological (body length and sex) information can assist in the interpretation of close kinship categories. Triad relationships in this figure result from one individual

occurring in more than one dyad. Dyads encircled by a dashed line show possible FS pairs. There were two triads comprising only female animals, and there were three triads comprising two males and one female.

The triad 93IVR003 (male), 05IVR45 (male) and 05IVR44 (female) was one of the strongly supported triads and also a case in which the biological data contributed to the interpreting of close kin inference. Male 93IVR003 had a dyad with a male 05IVR45 and another dyad with female 05IVR44. The two pairwise LOD scores were $LOD_{p0}=7.73$ and $LOD_{p0}=8.69$, respectively. The fact that 93IVR003 had reached 13.7 m in body length in 1993 indicates that it was sexually mature. This supports the possibility that the other two animals in the triad (05IVR45 and 05IVR44) are the son and daughter of male 93IVR003, respectively.

Geographical connections for inferred dyads

Figure 5 shows the geographical locations and connections for the 28 dyads with $LOD_{p0}>6$. Six of these dyads

were individuals sampled at low and high latitudes; the high-latitude individuals of these dyads were mostly concentrated between 85°E and 110°E. Within the Antarctic, parent-offspring relationships were concentrated within the Antarctic Indian sector. There was a single connection between the inner and outer Indo-Pacific sectors for one possible full-sibling relationship. In addition, there were no close kin relationships identified between individuals distributed west and east of 85°E.

Interpretation of results

Close kinship information from this study strongly supported the view that whales in the Indian sector of the Antarctic belong to a stock related to Southwestern Australia, as most of the dyads were found within the sector 85°–135°E (within and between austral summer seasons), and between this Antarctic sector and whales in lower latitudes off Southwestern Australia (between different austral summer seasons). Also, results of this study supported the hypothesis that fidelity to feeding areas is

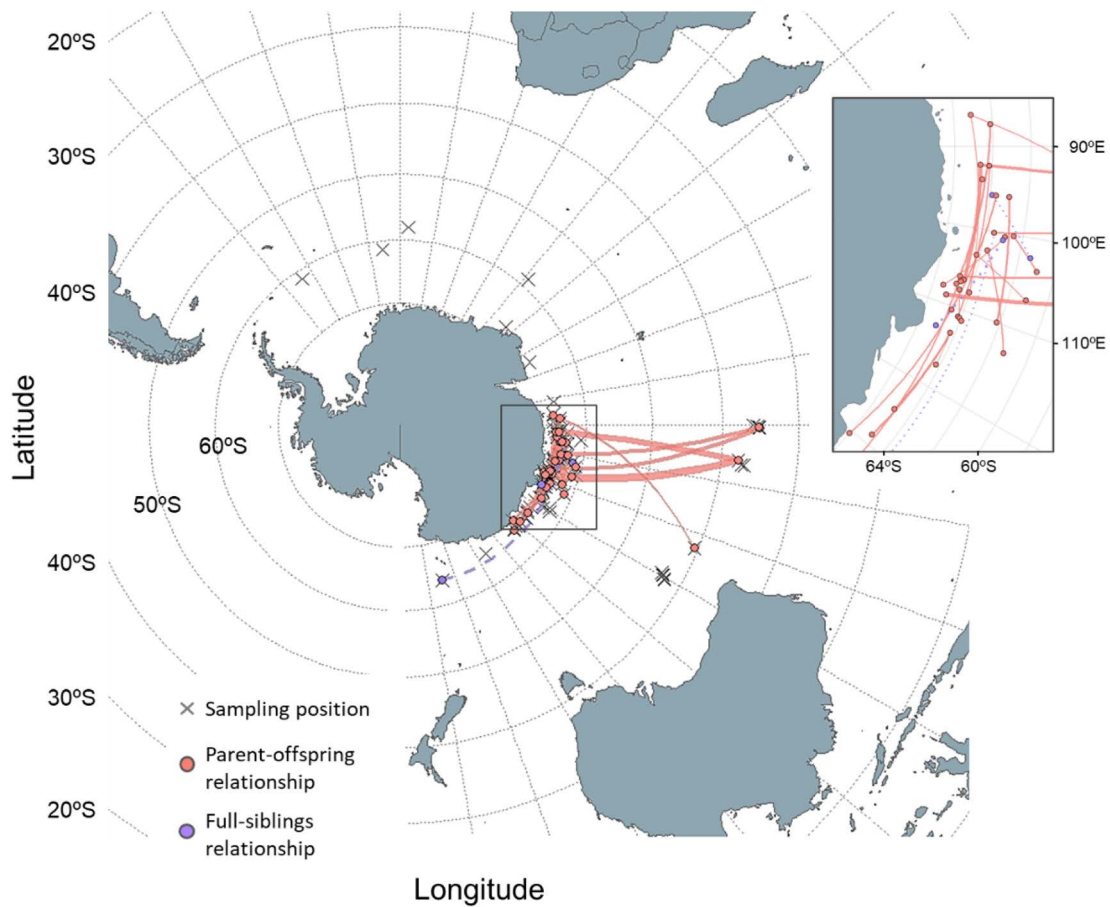


Figure 5. Geographical connections for dyads with $LOD_{p0}>6$ indicating the degree of strength of the connection. The inset shows a map focusing on the Indian sector of the Antarctic where most of the observations are concentrated. Maps were obtained by R (R core Team, 2023) using 'rnatuarearth' package (Massicotte and South, 2023).

inherited from mother to offspring (Carroll *et al.*, 2015; Valenzuela *et al.*, 2009), given that most of the maternally related parent offspring pairs occurred in a similar sector in the Antarctic (85°–135°E) or represented connections between low and high latitude waters in this longitudinal sector.

As previously mentioned, however, a possible full-sibling pair was found between the Antarctic Indian sector and the outer Amantarctic Indian sector (between 160° and 170°E). This could be explained by the sporadic dispersion of close kin whales of the Southwest Australia stock in the core longitudinal sector (85°–135°E) into a more eastern longitude. An alternative explanation could be that whales from different breeding grounds, for example, Southwest Australia and New Zealand, breed (Patenaude *et al.*, 2007) and their ‘hybrid’ offspring migrate to different sectors of the Antarctic. However, currently there is no evidence of monogamy, so full-siblings can be expected to be rare. A third interpretation is the possibility of a false-positive dyad for this particular case. The first interpretation of sporadic longitudinal dispersion of possible full siblings is considered more plausible.

The uncertainties for the inference of close kin dyads are related mainly to the amount of information in the DNA profiles, and also to genotyping error. In the present study, a simple error model with a constant and independent per-allele error rate was based on genotyping error information from the North Pacific right whale. Uncertainty associated with the analytical approach used in this study could be reduced by the use of other genetic markers, as well as biological information such as age of the individuals.

POTENTIAL USE OF CLOSE KINSHIP INFORMATION

Close kinship data from this study can be used for estimating the abundance of the stock using close-kin mark-recapture (CKMR) method (Bravington *et al.*, 2016). The CKMR method is an extension of classical mark-recapture methods. In previous applications of this method to fish species, age data have been available, which simplified the calculation of kinship recapture probabilities. No age data is available for southern right whales. Therefore if the CKMR method is applied to this species in the future, some modification of the method that addresses the lack of age data, will be required (e.g., Skaug, 2001).

As mentioned earlier, close kin information are potentially useful for estimating some demographic parameters, which should be monitored for conservation purposes. Furthermore, the information on close kinship

in this study could also be informative about inbreeding in this species.

However, when using feeding ground data, if the aim is to estimate the size of a breeding population, then the question of which population size is being estimated requires careful consideration, and further methodological developments are also needed.

See Takahashi *et al.* (2024) for details of the close kinship inference using LOD score, and application and results for southern right whales.

ACKNOWLEDGEMENTS

The document Takahashi *et al.* (2024) resulted from a research collaboration from ICR and the University of Bergen, Norway, represented by Prof. H.J. Skaug. We thank B. Følund, University of Bergen, for his assistance in the application of the analytical approach. This study was based mainly on genetic data published previously by Pastene *et al.* (2021) as well on new genetic data obtained in more recent years. The data were obtained from biopsy samples collected in different dedicated sighting surveys. We thank cruise leaders, captains, crew members, and researchers who participated in Japanese and IWC-sponsored surveys in the Antarctic for their effort in collecting biopsy samples and ancillary data of the southern right whales. We also thank M. Taguchi and other members of the ICR genetic team for producing the genetic data used in this study. Finally we thank the TEREP editorial team for their editorial check.

REFERENCES

- Abe, H., Goto, M., Pastene, L.A., Dewa, K. and Naito, E. 2001. Practical use of multiplex fluorescent PCR for Cetacean sex identification. *Mar. Mam. Sci.* 17: 657–664.
- Bannister, J. 2001. Status of southern right whales (*Eubalaena australis*) off Australia. *J. Cetacean Res. Manage.* 3: 103–110.
- Bannister, J.L., Pastene, L.A. and Burnell, S.R. 1999. First record of movement of a southern right whale (*Eubalaena australis*) between warm water breeding grounds and the Antarctic Ocean, South of 60°S. *Mar. Mam. Sci.* 15(4): 1337–1342.
- Bérubé, M., Jørgensen, H., McEwing, R. and Palsbøll, P.J. 2000. Polymorphic di-nucleotide microsatellite loci isolated from the humpback whale, *Megaptera novaeangliae*. *Mol. Ecol.* 9(12): 2181–2183.
- Buchanan, F.C., Friesen, M.K., Littlejohn, R.P. and Clayton, J.W. 1996. Microsatellites from the beluga whale *Delphinapterus leucas*. *Mol. Ecol.* 5(4): 571–575.
- Bravington, M.V., Skaug, H.J. and Anderson, E.C. 2016. Close-Kin Mark-Recapture. *Statistical Science* 31(2): 259–274.
- Carroll, E.L., Rayment, W.J., Alexander, A.M., Baker C.S., Patenaude, N.J., Steel, D., Constantine, R., Cole, R., Boren, L.J.

- and Childerhouse, S. 2014. Reestablishment of former wintering grounds by New Zealand southern right whales. *Mar. Mam. Sci.* 30(1): 206–220.
- Carroll, E.L., Baker, C.S., Watson, M., Alderman, R., Bannister, J., Gaggiotti, O.E., Gröcke, D.R., Patenaude, N. and Harcourt, R. 2015. Cultural traditions across a migratory network shape the genetic structure of southern right whales around Australia and New Zealand. *Sci. Rep.* 5: 16182.
- Carroll, E.L., Fewster, R.M., Childerhouse, S.J., Patenaude, N.J., Boren, L. and Baker, C.S. 2016. First direct evidence for natal wintering ground fidelity and estimate of juvenile survival in the New Zealand southern right whale *Eubalaena australis*. *PLoS One.* 11(1): e0146590.
- Cranswick, A. 2022. New Zealand Southern Right Whale *Eubalaena australis*—Foraging Ecology and Distribution. The University of Auckland, 2022, Master thesis.
- Frasier, T., Rastogi, T., Brown, M., Hamilton, P., Kraus, S. and White, B. 2006. Characterization of tetranucleotide microsatellite loci and development and validation of multiplex reactions for the study of right whale species (genus *Eubalaena*). *Mol. Ecol. Notes* 6: 1025–1029.
- International Whaling Commission. 2001. Report of the workshop on the comprehensive assessment of right whales: A worldwide comparison. *J. Cetacean Res. Manage.* (Suppl.) 2: 1–35.
- International Whaling Commission. 2013. Report of the IWC workshop on the assessment of southern right whales. *J. Cetacean Res. Manage.* (Suppl.) 14: 437–462.
- Massicotte, P., South A. and Hufkens, K. 2023. *rnaturalearth*: World Map Data from Natural Earth. <https://docs.ropensci.org/rnaturalearth/>, <https://github.com/ropensci/rnaturalearth>.
- Palsbøll, P.J., Bérubé, M., Larsen, A.H. and Jorgensen, H. 1997. Primers for the amplification of tri- and tetramer microsatellite loci in baleen whales. *Mol. Ecol.* 6(9): 893–895.
- Pastene, L.A., Goto, M., Acuña, P., Taguchi, M., Hakamada, T. and Matsuoka, K. 2018. Population identity, site-fidelity, movement ranges and preliminary estimates of abundance of southern right whales in the Antarctic Indian sector inferred from genetic markers. Paper WG-EMM-18/18 presented to the CCAMLR WG-EMM, July 2018 (unpublished). 13 pp. [Available from <https://meetings.ccamlr.org/en/wg-emm-18>].
- Pastene, L.A., Goto, M., Taguchi, M. and Matsuoka, K. 2021. Genetic matches of southern right whales in the Indian sector of the Antarctic: a contribution towards understanding their movement and site-fidelity. *Cetacean Population Studies* 3: 129–138.
- Pastene, L.A., Taguchi, M., Lang, A., Goto, M. and Matsuoka, K. 2022. Population genetic structure of North Pacific right whales. *Mar. Mam. Sci.* 38(3): 1249–1261.
- Patenaude, N.J., Portway, V.A., Schaeff, C.M., Bannister, J.L., Best, P.B., Payne, R.S., Rowntree, V.J., Rivarola, M. and Baker, C.S. 2007. Mitochondrial DNA diversity and population structure among southern right whales (*Eubalaena australis*). *J. Hered.* 98(2): 147–157.
- R Core Team. 2023. R: A Language and Environment for Statistical Computing Austria, Vienna: R Foundation for Statistical Computing. Available from: <https://www.R-project.org/>.
- Skaug, H.J. 2001. Allele-sharing methods for estimation of population size, *Biometrics* 57(3): 750–756.
- Skaug, H.J., Bérubé, M. and Palsbøll, P.J. 2010. Detecting dyads of related individuals in large collections of DNA-profiles by controlling the false discovery rate. *Mol. Ecol. Resour.* 10(4): 693–700.
- Taguchi, M. 2020. Genetic tagging technique: basic concept and a case study by the Institute of Cetacean Research. *Technical Reports of the Institute of Cetacean Research (TEREP-ICR)* No.4: 35–40.
- Takahashi, M., Føland, B., Pastene, L.A. and Skaug, H.J. 2024. Geographical distribution of close kin in southern right whales on feeding grounds. *PLoS One.* <https://doi.org/10.1371/journal.pone.0301588>.
- Valenzuela, L.O., Sironi, M., Rowntree, V.J. and Seger, J. 2009. Isotopic and genetic evidence for culturally inherited site fidelity to feeding grounds in southern right whales (*Eubalaena australis*). *Mol. Ecol.* 18(5): 782–791.
- Valsecchi, E. and Amos, W. 1996. Microsatellite markers for the study of cetacean populations. *Mol. Ecol.* 5(1): 151–156.