Finding parent-offspring pairs among western North Pacific common minke whales

Tiedemann R, Tiedemann MR, Goto M, Taguchi M, Pastene LA

INTERNATIONAL
WHALING COMMISSION

Finding parent-offspring pairs among western North Pacific common minke whales

Ralph Tiedemann[1], Magnús R. Tiedemann[2], Mutsuo Goto[3], Mioko Taguchi[3] and Luis A. Pastene[3]

[1]*Unit of Evolutionary Biology/Systematic Zoology, Institute of Biochemistry and Biology, University of Potsdam, D-14476 Potsdam, Germany*

[2]*Faculty of Computer Science, University of Magdeburg, D-39016 Magdeburg, Germany*

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

ABSTRACT

A dataset of complete genotypes at 16 microsatellite loci, accompanied with mtDNA and biological information, in 4,554 North Pacific common minke whales was used to infer Parent-Offspring (P-O) relationships, using a Maximum-Likelihood approach. The relationship between False Discovery Rate (FDR) and Power (P) was evaluated by simulation. Of 145 inferred P-O pairs at an estimated FDR of 0.1, 141 were further evaluated by typing 10 additional microsatellite loci. 75 were confirmed (among them 26 Mother-Fetus pairs), 66 pairs were ranked "False Positives", yielding an overall observed FDR of 0.468. FDR_O was substantially reduced when J and O stock were analysed separately. While observed and estimated values for Power were in the same range of magnitude, observed FDR was always substantially higher than estimated FDR. This was attributed to the fact that FDR_E was estimated via simulation, implicitly assuming a single panmictic population, an assumption clearly not met in the present data set. This interpretation is corroborated by the reduced FDR_O when stocks were analysed separately. The dataset with 26 microsatellites clearly outperformed the 16 microsatellite data sets. At FDR_E=0.001, Power was at or close to 100% (P_E=0.989 and P_O=1.000) and the observed False Discovery Rate was FDR_O=0.128. Among the validated P-O pairs, O stock pairs were significantly overrepresented, while pairs between J and O stock individuals were absent. Specimens neither assigned to J nor O stock ("unassigned") exhibited a stronger affinity to the O stock. The J stock seems to appear on both sides of Japan closer to the coast, while the O stock occurs mostly east of Japan, both close to the coast and far offshore. This analysis provides no evidence for further stock structure in the area covered by this data set. This study demonstrates that a modest increase in the number of loci investigated (here, from 16 to 26 microsatellite loci) may already substantially improve kinship inference under Maximum Likelihood. It further addresses recommendations made at both the JARPNII final review and the 2016 IWC Scientific Committee meeting regarding kinship analysis in North Pacific common minke whales.

INTRODUCTION

Results of the updated analyses on stock structure of western North Pacific common minke whales were presented to the JARPNII final review workshop (IWC, 2017a), and a summary of the new information was made by Pastene (2017). Most of the new analyses were conducted in response to recommendations made during the 2009 JARPNII mid-term review (IWC, 2010).

One of the recommendations made by the mid-term review was the examination of 'the spatial distribution of related individual, which can provide information directly relevant to stock structure consideration' (IWC, 2010). Subsequently the analysis on close kin in western North Pacific common minke whale started in 2015 applying the likelihood-based approach described in Tiedemann *et al*. (2014), and preliminary results were presented during the open session of the JARPNII final review workshop.

In reviewing the preliminary results in 2016, the International Whaling Commission Scientific Committee (IWC SC) recommended that a paper to examine the spatial distribution of close kin in North Pacific

minke whales be submitted for review at the 2017 annual meeting. The IWC SC made the following specific suggestions:

a) Emphasises the importance of evaluating the potential for false positive and false negative detections of parent-offspring pairs (see Tiedemann *et al.*, 2014);
b) Encourages the authors to explore different approaches (e.g. software) to conduct kinship-based analyses; and
c) Recommends that the samples be genotyped at additional loci (microsatellites or SNPs to validate the putative Parent-Offspring pairs that were identified (IWC, 2017b).

The objective of the present study was to conduct a comprehensive analyses on kinship (Parent-Offspring pairs) based on i) the total available genetic samples till 2015 from JARPN, JARPNII and Japanese by-catches; ii) microsatellite profiles; iii) the analytical approach described in Tiedemann *et al.*, 2014; and iv) the recommendations from the IWC SC in 2016.

Furthermore, this paper evaluates the effect of implicit assumptions inherent to LOD-score-based kinship analysis. Specifically, population simulations to estimate False Discovery Rate (FDR_E) and Detection Power (P_E) based on observed genotype frequencies may implicitly assume random mating and lack of population/stock structure, while real data sets frequently do not meet this assumption. Therefore, we also evaluate the impact of data partitioning on kinship inference.

MATERIALS AND METHOD

Genetic and Biological data
Offshore samples of common minke whales from the western North Pacific were from JARPN/JARPNII surveys from 1994 to 2013 at SA7, SA8, SA9, and SA11. Common minke whale samples obtained from the coastal JARPNII survey between 2002 and 2015 were also used in this study, Kushiro in SA7CN and Sanriku in SA7CS. Samples from common minke whales that were bycaught on set net fishery along the Japanese coast from 2001 to 2015 were also used. The by-catches used were from subareas SA2, SA6, SA7, SA10, and SA11 year-round.

This analysis was based on a dataset of a total of 4,554 North Pacific minke whales including fetuses for which information on 16 microsatellite loci (Pastene *et al.*, 2016), mtDNA haplotype (M. Goto, unpubl. data), and biological information (e.g. sampling date and position, sex, sexual maturity and body length) was available at ICR Japan. The entire data set was analysed for occurrence of Parent-Offspring (P-O) pairs according to Tiedemann *et al.* (2014) (see below). For all P-O pairs inferred at an estimated False Discovery Rate (FDR_E) of 0.1, the respective specimens (n=275) were typed at an additional set of 10 microsatellites (GATA97408, GATA91083, GATA43950, GATA38314, GATA6059012, GATA5946992, GATA6063862, GATA6064765, GATA5943219 and GATA5890064: P. J. Palsbøll, pers. comm.). For four individuals, this typing was not complete, such that this analysis yielded full genotypes at in total 26 loci for 271 individuals.

Inference of Parent-Offspring Pairs
From the available genotype data for the 16 standard microsatellite loci, we formed five different data sets, i.e., the full data set, the "J stock only" dataset, the "O stock only" dataset, the "J stock+unassigned" dataset, and the "O stock+unassigned" dataset, based on stock assignment in Pastene *et al.* (2016).

The five data sets were analyzed for occurrence of Parent-Offspring (P-O) pairs according to Tiedemann *et al.* (2014) with the following parameters:
- Typing error rate per allele 0.0025 (Goto *et al.* 2016)
- Number of random individuals simulated 5000 (corresponding to 12,497,500 random pairs for estimation of the False Discovery Rate (FDR_E)
- Number of related pairs simulated 100,000 for estimation of Power P_E

Only specimens with complete genotypes were included in the analyses and fetuses were excluded from the population-wise allele-frequency calculation (see Tiedemann *et al.* 2014 for details).

We ran the analysis with the following FDR_E-values: 0.3; 0.2; 0.1; 0.05; 0.01; 0.001 to evaluate the impact of FDR_E on Power. P-O pairs inferred at an FDR_E of 0.1 were retained for further validation.

Validation of Parent-Offspring Pairs
Validation of Parent-Offspring pairs was performed in two steps as follows:

First, the results at the 10 additional microsatellite loci were considered:
- For the dataset of all specimens involved in an inferred P-O pair based on 16 loci (FDR 0.1), the analysis was repeated with 26 microsatellites, using again FDR-values of 0.3; 0.2; 0.1; 0.05; 0.01; and 0.001. Note that this data set allowed only for 36,585 pairwise comparisons. Further, it is not a random data set, as the likelihood to be part of a P-O relationship is elevated: as any specimen was part of a inferred P-O relationship, it could be as high as 0.000027 (if all inferred P-O pairs were true positives). If correcting this rate for the observed False Discovery Rate (FDR_O=0.468 for the full data set at FDR_E 0.1; see table 4), the rate of P-O pairs in this dataset is still 0.000015, about 2 times higher than in the full data set.
- All inferred P-O pairs were checked for full compatibility with a P-O relationship at the 10 additional microsatellite loci, i.e., sharing at least one allele at any locus. Note that this implicitely assumes no genotyping error and no mutation. It may hence produce False Negatives, i.e., rejection of valid P-O pairs.
- Mother-Fetus pairs were considered as *biologically validated* true positives. Out of 26 Mother-Fetus pairs among the inferred P-O pairs, 24 were fully compatible with a P-O relationship at the 10 additional microsatellite loci, but two (8%) were not (cf. "no" for full fit in table 2), due to a single mismatch each. The lowest LOD score for a Mother-Fetus pair in the 26 loci microsatellite analysis was 4.11 – this pair was one of the two with a "no" for full fit (Table 2).
- All other inferred P-O pairs with a "no" and a LOD score below 4.11 were considered False Positives (FP).
- 85 P-O pairs yielded a LOD-Score of at least 4.11, i.e., all 26 Mother-Fetus pairs and 59 other inferred P-O pairs. Among these 59 pairs, 49 had a "yes" for full fit at the additional 10 microsatellite loci and were retained. 10 (17%) had a "no" here, which is about twice the percentage for such cases when compared to the Mother-Fetus pairs (8% with a "no" for full fit). We ranked those 10 pairs by LOD score and assigned the five pairs with the lowest LOD scores as "false positives". The other 5 pairs (all with a LOD score over 5.1) comprised one Father-Son pair (retained) and 4 Mother-Offspring pairs. These 4 pairs were evaluated for their mtDNA haplotype: 3 pairs exhibited a shared type (retained), while one pair had different mtDNA types (ranked False Positive; see Table 2 for details).

All inferred P-O pairs were independently checked for compatibility of the inference (true vs. false positive) against the biological data, i.e., mtDNA in Mother-Offspring pairs and compatibility of year of catch and sexual maturity, i.e., at least one sexually mature adult in the pair.

- Out of the 63 pairs ranked False Positives by the microsatellite genetic inference (26 loci), 34 had been independently ranked "False Positives" by biological information incompatible with a P-O relationship, i.e., both specimens immature or inferred Mother-offspring pair with different mtDNA haplotype.
- Out of the 78 pairs retained after microsatellite genetic inference, 3 were Mother-Offspring pairs with incompatible mtDNA haplotype. They were additionally considered "False Positives", rendering a set of 75 pairs considered "True positives" (i.e., real P-O pairs). Only three of them (all bycatch) had some ambiguity with regard to whether their length measurement is fully compatible with adulthood. As length measurements during bycatch-operations may be less precise and as two of these three pairs were Mother-Offspring pairs supported by shared mtDNA type, these three pairs were retained as "True Positives" (see Table 2 for details).

Calculation of observed False Discovery Rate and Power
Observed False Discovery Rate (FDR_O) was calculated for estimated FDR_E rates of 0.1, 0.05, 0.01, and 0.001 for all datasets as the rate of inferred P-O pairs ranked "False Positives" after validation (see above).

Observed Power (P_O) was calculated for estimated FDR_E rates of 0.1, 0.05, 0.01, and 0.001 for all datasets as the percentage of Mother-Fetus pairs in the data set which were detected by the analysis.

RESULTS

In total, we inferred – based on an FDR of 0.1 and 16 microsatellite loci - 145 Parent-Offspring pairs, thereof 27 mother-fetus pairs. Genotyping additional 10 microsatellite loci of the involved 275 individuals yielded full genotypes for 271 individuals. For 4 pairs (thereof 1 mother-fetus pair), genotypes were incomplete, such that 141 P-O pairs (thereof 26 mother-fetus pairs) were retained for validation (Tables 1, 2).

Of the 26 Mother-Fetus pairs, 24 were genetically fully compatible with a P-O relationship (i.e., one shared allele at each of the additional 10 microsatellite loci), while in 2 pairs, there was a single mismatch at one of the loci. If we assume no mutation, this translates into a typing error rate of 0.0038 per allele, close to the error rate of 0.0025 previously estimated for the standard 16 microsatellite loci in the entire NP minke whale data set (Goto *et al.*, 2016).

Out of the 141 P-O pairs with full genotype at 26 microsatellite loci, 75 P-O pairs were validated as "True Positives": 26 were Mother-Fetus pairs (i.e., considered *biologically validated* True Positives), the remaining 49 were validated by the additional 10 microsatellite loci (Table 2; see methods for details). In all cases, biological information was compatible with a Parent-Offspring relationship (i.e., at least one specimen was an adult; time between catches was compatible with P-O relationship). 10 of these validated P-O pairs were Mother-Offspring pairs, all confirmed by identical mtDNA haplotype (Table 2).

66 P-O pairs inferred based on 16 microsatellites at an FDR_E of 0.1 were ranked "False Positives" after validation. For 31 of them, this judgement was confirmed by biological information incompatible with a P-O relationship, i.e., both specimens immature or inferred Mother-Offspring pair with different mtDNA haplotype.

The observed False Discovery Rate (FDR_O) of 0.468 was strongly biased towards J-stock individuals, i.e., only 9 out of 46 inferred P-O pairs originating from J stock were validated ($FDR_O=0.804$), while for the O-stock, 57 out of 83 were validated ($FDR_O=0.313$). If stocks were analysed separately, FDR_O dropped dramatically in the J stock (i.e., to 0.000) and also substantially in the O stock (to 0.258; Table 3).

If FDR_E was decreased in the initial analysis on the full data set (16 microsatellites) to 0.01, observed FDR_O was also considerably lower, i.e., 0.179 for the full data set, 0.000 for the J-stock, and 0.103 for the O-stock (Table 3).

Observed Power (P_O) was in the same order of magnitude as estimated power (P_E). For the full data set, P_O was slightly higher than P_E, in the data subsets involving the O-stock, this relation was reversed (Table 3). In the data partitions "J stock only" and "J+unassigned", there were only very few Mother-Fetus pairs (2 and 4, respectively; see Table 1). As none of these Mother-Fetus pairs was detected in the Maximum Likelihood analysis, this nominally translates into P_O of 0.000, but this estimate is likely imprecise and should be treated with caution.

With an overall $P_O=0.5$ for the full data set at $FDR_E=0.1$ (the value used for the inference presented in Table 1), we can expect to have identified about half of the true P-O pairs present in the data set.

For the dataset with 26 microsatellites, both estimated and observed Power were considerably higher: At $FDR_E=0.001$, $P_E=0.989$ and $P_O=1.000$. Under these conditions, observed False Discovery Rate was $FDR_O=0.128$.

If we evaluate how validated P-O pairs were distributed across stocks, there were 3 significant deviations from random expectations (Table 4):

- There were significantly more observed P-O pairs in the O stock than expected.

- There were significantly fewer (i.e., no) observed P-O pairs among the J stock and the O stock than expected under a random association.
- There were significantly fewer (i.e., no) observed P-O pairs among the J stock and unassigned specimens than expected under a random association.

Geographically, many validated P-O pairs were identified across subareas, with J stock pairs centered around the coast at both sides of Japan, while the O stock pairs appeared mostly east of Japan and several validated O stock P-O pairs connected subareas close to Japan with areas far offshore (Figure 1).

DISCUSSION

Using kinship analysis to inform stock structure discussions has been repeatedly advocated when large genetic data sets are at hand (e.g., Skaug *et al*., 2010; IWC, 2010; IWC, 2017b). Inference is often based on assigning the likelihood of a given pairwise relationship (most frequently, Parent-Offspring) relationship to pairs of individuals, based on genotypes and population-wide allele frequencies (Skaug *et al*., 2010; Benónísdóttir, 2012; Tiedemann *et al.*, 2014).

Given a particular data set, there is a trade-off between the False Discovery Rate (FDR) and the Power (P) to detect dyads of related individuals (Skaug *et al*., 2010; Benónísdóttir, 2012; Tiedemann *et al*., 2014). Both FDR and P are typically estimated by simulation based on population-wide allele frequencies (see, e.g., Tiedemann *et al*. 2014). However, such simulations implicitly assume random mating and lack of population/stock structure, while real data sets frequently do not meet this assumption.

In most cases, estimated values of FDR and P (hereafter, FDR_E and P_E) are not explicitly validated. Here, we use the rather unique data set of North Pacific common minke whales (4,554 specimens with full genotype information at 16 loci, mitochondrial DNA, and associated biological information) for a validation of these values. Specifically, Parent-Offspring (P-O pairs) inferred at a given estimated False Discovery Rate (FDR_E) were genotyped at 10 additional microsatellite loci for validation, both by likelihood inference and locus-wise inspection for fit with a P-0 relationship. Furthermore, biological information (length, sexual maturity, mtDNA type) was utilized to confirm or contradict inferred P-O pairs in order to arrive at validated "True Positives" of P-O relationships, as well as an observed value for the False Discovery Rate (FDR_O).

We utilized 53 Mother/Fetus pairs in the data set as *biologically validated* P-O pairs. Observed Power (P_O) could be calculated as the proportion of these Mother/Fetus pairs detected under the various False Discovery Rates and data partitions applied.

Observed vs. Estimated values of False Discovery Rate and Power
Our analysis clearly shows that estimations of False Discovery Rates based on random individuals substantially underestimates the rate by which False Positives are actually detected (i.e., $FDR_E \ll FDR_O$). This is perhaps not surprising when considering that real populations do not comprise random assemblages of alleles, but are constituted of individuals many of which related to one another to various degrees. It remains to be inferred whether "False Positives" for P-O relationship may constitute, at least in part, relatives of a less close relationship (e.g., half sibs or cousins). This pattern of "False Discovery" is particularly apparent, if specimens of different populations/stocks are lumped into a single data set and subsequently treated as if they would constitute a single panmictic population (here, J and O stock). Indeed in our analysis, FDR_O was substantially reduced, if stocks were analyzed separately.

This effect can be asymmetric, as – in our analysis – there was a much higher FDR_O for the J than for the O stock, when analyzing them jointly. The O stock is genetically considerably more diverse than the J stock (e.g., average microsatellite allele number per locus: 11.94 vs. 8.44) and less inbred (inbreeding coefficient F_{IS} 0.002 vs. 0.008) (Goto *et al*., in prep). Apparently, the (relatively larger) genetic homogeneity in the J stock leads to an increased rate of erroneously inferred P-O pairs, when jointly analysed with the more diverse O stock. Interestingly, FDR_O drops from 0.804 (combined data set) to 0.000, when the J stock is analysed separately.

Observed Power (P_O) was in most data partitions lower than estimated Power (P_E) (except for the full data set), but the differences were not as large as for the FDR: Generally, P_O and P_E values were in the same order of magnitude. It remains to be further evaluated whether nominal differences between P_O and P_E just constitute stochastic variation or whether there is any systematic bias in these estimates.

The dataset with 26 microsatellites clearly outperformed the 16 microsatellite data set. Even at $FDR_E=0.001$, Power was very high ($P_E=0.989$ and $P_O=1.000$) and observed False Discovery Rate was relatively low ($FDR_O=0.128$). Because this data set is considerably smaller (271 vs. 4,554 specimens) and has a doubled internal likelihood for occurrence of P-O pairs, these measures may contain some bias and are likely not as precise as those for the full data set. Nonetheless, these data provide strong experimental support for the benefit of additional loci in kinship analyses.

Stock affinity and geographical position of validated P-O pairs
Not surprisingly, the data strongly support the distinctiveness of the J and the O stock in North Pacific common minke whales. It should be noted that – for the 16 microsatellite data set – this is not independent evidence, as the same 16 microsatellites had been utilized to infer the stock assignment (Pastene *et al.,* 2016). However, the addition of 10 further microsatellite loci did not change this pattern.

In essence, the J stock appears to be genetically distinct with lower variability and a geographic restriction to the waters around the coasts of Japan. The O stock is genetically more diverse and geographically widely distributed from the East coast of Japan to far offshore. The cohort of "unassigned" specimens shows a significant affinity to the O stock: P-O pairs across O stock appeared as often as expected under the assumption of a single population, while there were no validated P-O pairs across J stock and "unassigned" specimens.

Conclusive remarks
The large North Pacific common minke whale genetic data set, the available biological information, the abundance of Mother-Fetus pairs in the data set, and the additional typing of 10 microsatellites for inferred Parent-Offspring pairs offered the rare opportunity of a validation of kinship inferences and a calculation of observed values for False Discovery Rate and Power, as has been recommended in the JARPNII final review (IWC 2017b) and the IWC Scientific Committee. This analysis shows that:

- With the standard set of 16 microsatellites, high observed power is associated with quite high observed False Discovery Rates.
- Joint analysis of different stocks considerably increases the False Discovery Rate, in particular if one of the stocks is genetically less variable/more inbred.
- With the addition of 10 additional microsatellite loci, an observed Power of 1.000 could be combined with an observed False Discovery Rate of around 0.1.

If applied for other studies or other microsatellite loci, these measures may evidently vary depending on stock structure and locus characteristics. Nonetheless, we argue that our study provides experimental evidence that a modest increase in the number of loci investigated (here, from 16 to 26 loci) may already substantially improve kinship inference under Maximum Likelihood.

With regard to stock structure in North Pacific common minke whales, the distinctiveness of J and O stock is clearly seen in the analysis. Hitherto "unassigned" specimens show more affinity to the O stock. The J stock seems to appear on both sides of Japan closer to the coast, while the O stock occurs mostly east of Japan, both close to the coast and far offshore. This analysis provides no evidence for further stock structure in the area covered by this data set.

REFERENCES

Benónísdóttir S (2012) Statistical approach to relatedness analysis in large collections of genetic profiles – an application to a DNA registry of Fin whales. M.Sc. thesis, University of Bergen.

Goto, M., Yoshida, H. and Pastene, L.A. 2016. A note with an estimate of microsatellite genotyping error rate for common minke whales. Paper SC/66b/DNA/01 presented to the IWC Scientific Committee, June 2016 (unpublished). 3pp.

International Whaling Commission. 2010. Report of the expert workshop to review the ongoing JARPNII programme. *J. Cetacean Res. Manage*. 11 (Suppl. 2): 405-449.

International Whaling Commission. 2017a. Report of the Expert Panel of the final review on the western North Pacific Japanese Special Permit programme (JARPNII). *J. Cetacean Res. Manage.* 18 (Suppl.): *in press.*

International Whaling Commission. 2017b. Report of the Scientific Committee. *J. Cetacean Res. Manage.* 18 (Suppl.): *in press.*

Pastene, L.A., Goto, M., Taguchi, M. and Kitakado, T. 2016. Temporal and spatial distribution of the 'J' and 'O' stocks of common minke whale in waters around Japan based on microsatellite DNA. Paper SC/F16/JR38 presented to the IWC Scientific Committee, February 2016 (unpublished). 14pp.

Pastene, L.A. 2017. Annex I, Appendix 3. A summary of the new information on stock structure in the western North Pacific common minke whale presented to the JARPNII final review. *J. Cetacean Res. Manage.* 18 (Suppl.): *in press.*

Skaug HJ, Bérubé M, Palsbøll P (2010) Detecting dyads of related individuals in large collections of DNA-profiles by controlling the false discovery rate. *Molecular Ecology Resources* 10, 693-700.

Tiedemann, R., Tiedemann, M.R., Gunnlaugsson, D., Pampoulie, C. and Vikingsson, G.A. 2014. Finding relatives among North Atlantic common minke whales (*Balaenoptera acutorostrata*) based on microsatellite data: the relationship between false discovery rate (FDR) and detection power. Paper SC/65b/RMP05 presented to the IWC Scientific Committee, May 2014 (unpublished). 8pp.

Table 1: Sample details for used data sets. P-O inferences are from the full data set (in parentheses inferences from separate stock-specific data sets). Stock assignment according to Pastene *et al.* (2016)

	Full dataset	Only J stock	Only O stock	J stock + unassigned	O stock + unassigned
n	4582	1779	2376	2206	2803
- thereof fetuses	54	2	50	4	52
n with full genotypes for 16 loci	4554	1767	2366	2188	2787
- thereof fetuses	53	2	49	4	51
% fetuses in dataset	0.012	0.001	0.021	0.002	0.018
Tested pairs	10367181	1560261	2797795	2392578	3882291
Inferred P-O pairs at FDR 0.1	145	47 (5)	86 (67)	48 (7)	97 (81)
- thereof full genotypes at 26 loci	141	46 (4)	83 (62)	47 (5)	94 (72)
number of confirmed P-O pairs	75	9 (4)	57 (46)	9 (4)	66 (55)
% confirmed P-O pairs of all pairs	0.00000723	0.00000577	0.00002037	0.00000376	0.00001700
Detected Mother-Fetus pairs at FDR 0.1	27	0 (0)	23 (17)	0 (0)	27 (21)
- thereof full genotypes at 26 loci	26	0 (0)	22 (15)	0 (0)	26 (19)

Table 2: Inference of Parent-Offspring (P-O) relationships. P-O Inferences of the full data set (16 microsatellite loci; n=4,554; FDR≤0.1) were evaluated by the following means: All specimens of the inferred pairs were typed at 10 additional microsatellite loci; Inference of P-O pairs was repeated for this subset (26 microsatellites; n=271). Further, at the 10 additional loci, a fit with a P-O relationship (i.e., at least one matched allele at any locus) was evaluated. Full fit means that all 10 loci are compatible with a P-O relationship. Biological data was used to support (##matching mtDNA in Mother-Offspring pairs) or contradict (#mismatch in mtDNA in Mother-Offspring pairs or both individuals of immature status or length) P-O relationships (see methods for details). For confirmed P-O pairs, the type of relationship as well as the area of origin are provided.

FP=False Positive; FDR_E=False Discovery Rate; P_E=Power (estimated by simulations); (*) inferred at FDR≤0.1; * inferred at FDR≤0.05; ** inferred at FDR≤0.01; *** inferred at FDR≤0.001; ### multiple offspring inferred for same parent. Stock assignment: O=O stock; J=J stock; U=unassigned; cf. Pastene *et al.*, (2016)

		Stock	Full data set 16 Loci		Subset 26 Loci		Subset 10 Loci			
			LOD Threshold	P_E	LOD Threshold	Power				
FDR_E	0.3		3.673063241	0.796	-5.049196083	1.000				
FDR_E	0.2		4.331695124	0.617	-4.046957537	1.000				
FDR_E	0.1		4.921379986	0.437	-2.37039593	1.000				
FDR_E	0.05		5.112659013	0.383	-0.985059046	1.000				
FDR_E	0.01		5.690162171	0.240	1.347098689	1.000				
FDR_E	0.001		5.690162171	0.240	3.708856137	0.989				
Individual_1 (I_1)	Individual_2 (I_2)	I_1I_2	LOD Value		LOD Value		Full fit	Interpretation	Subarea I_1	Subarea I_2
NP94MI017	NP05MI019	OO	4.96150295	(*)	-2.539504471	ns	no	FP		
NP94MI019	NP01MI079	OO	4.990618879	(*)	6.231414457	***	yes	Father-Son	9	8
NP95MI025	NP96MI067	OU	5.388497561	*	6.590064324	***	yes	Father-Son	9	7CN
NP95MI032	NP99MI004	OO	5.224221146	*	-0.51393265	*	no	FP		
NP96MI051	CS08MI017	OO	5.55281185	*	-7.432857429	ns	no	FP#		
NP97MI002	NP01MI052	OO	6.2682621	***	8.985446939	***	yes	Daughter-Father	9	9
NP97MI028	CK05MI042	OO	5.238267556	*	-0.276541245	*	no	FP#		
NP97MI036	NP08MI011	OO	6.361714917	***	2.583408749	**	no	FP		
NP97MI041	NP98MI050	OO	5.420573721	*	6.543548357	***	yes	Father-Son	9	7E
NP97MI059	NP98MI040	OO	5.153040633	*	-0.012606712	*	no	FP		
NP98MI029	CK14MI011	OO	7.103259317	***	9.316438867	***	yes	Son-Father	7WR	7CN
NP98MI043	NP00MI016	OO	7.227285909	***	8.428513014	***	yes	Father-Son	7E	9
NP99MI013	NP02MI004	OO	5.047657352	(*)	-2.706846076	ns	no	FP		

NP99MI013	14BC058	OO	5.18774905	*	-3.159217003	ns	no	FP#		
NP99MI018	NP02MI030	OO	5.05663181	(*)	7.487393501	***	yes	Father-Son	7CN	8
NP99MI023	NP06MI018	OO	4.985257975	(*)	0.354713363	*	no	FP#		
NP99MI032	NP10MI009	OO	5.078096243	(*)	-13.56208246	ns	no	FP		
NP99MI093	14BC030	JJ	5.772674741	***	5.019246683	***	no	FP		
NP99MI098	03BC088	JJ	5.301571497	*	1.849251568	**	no	FP		
NP00MI030	NP12MI025	OO	6.376023688	***	-0.802177126	*	no	FP#		
NP01MI034	CK09MI003	UO	4.978835842	(*)	-1.570407376	(*)	no	FP#		
NP01MI092	14BC054	OO	6.733999537	***	8.397905477	***	yes	Son-Mother##	9	7CS
NP01MI094	CK04MI041	OO	5.008975645	(*)	-7.546925608	ns	no	FP		
NP01MI100	CK02MI035	OO	5.428208345	*	9.493898933	***	yes	Father-Daughter	9	7CN
NP02MI005	NP02MI016	OO	5.099718419	(*)	0.047976082	*	no	FP#		
NP02MI010	08BC065	OO	4.98034817	(*)	-4.954903156	ns	no	FP		
NP02MI028	NP04MI026	OO	6.966459578	***	9.38298552	***	yes	Father-Son	8	9
NP02MI060	03BC051	OO	5.269387838	*	7.888491571	***	yes	Son-Mother##	7CN	7CS
NP02MI079	NP10MI002	OO	4.952452294	(*)	-5.362461564	ns	no	FP		
NP02MI087	NP02MI087F	OO	6.363904788	***	10.49397717	***	yes	Mother-Fetus	7CN	
NP03MI009	NP03MI009F	OO	6.29765356	***	9.156212598	***	yes	Mother-Fetus	7WR	
NP03MI011	NP07MI100	OU	5.403248091	*	-6.903202723	ns	no	FP		
NP03MI038	CK13MI039	OO	6.626260945	***	7.001652874	***	yes	Father-Son	8	7CN
NP03MI045	NP03MI045F	OO	5.040166394	(*)	9.441137662	***	yes	Mother-Fetus	8	
NP03MI054	NP03MI054F	UO	5.388751508	*	7.830643859	***	yes	Mother-Fetus	8	
NP03MI073	CS03MI012	OU	5.944545833	***	8.558190489	***	yes	Father-Son###	9	7CS
NP03MI073	CK15MI022	OO	6.50994481	***	7.695996221	***	yes	Father-Son###	9	7CN
NP03MI075	06BC073	OO	5.948027705	***	7.924329903	***	yes	Father-Son	9	7CS
NP03MI098	02BC052	OO	7.951766683	***	10.90533317	***	yes	Father-Daughter	9	7CS
NP03MI099	NP03MI099F	OO	6.070856519	***	7.722216955	***	yes	Mother-Fetus	9	
NP04MI025	CS11MI013	OO	6.068881686	***	8.281565364	***	yes	Father-Daughter	9	7CN
NP04MI029	CS06MI052	OO	5.536189841	*	-3.778967714	ns	no	FP		

NP04MI051	NP04MI051F	OO	5.823847066	***	8.742311144	***	yes	Mother-Fetus	9	
NP04MI068	NP04MI068F	OO	6.380774719	***	8.670532719	***	yes	Mother-Fetus	9	
NP04MI075	NP12MI013	OO	5.080019419	(*)	-2.226150339	(*)	no	FP#		
NP04MI077	NP07MI083	OO	7.507904896	***	8.546866706	***	no	Father-Son	9	7CN
NP05MI046	CK07MI029	OO	4.994978692	(*)	-10.21854182	ns	no	FP		
NP05MI084	CS07MI022	OO	8.216023522	***	1.417784428	**	no	FP#		
NP05MI084	NP05MI084F	OO	7.734071389	***	10.07059881	***	yes	Mother-Fetus	9	
NP06MI026	NP06MI026F	OO	5.290664836	*	7.611251886	***	yes	Mother-Fetus	7E	
NP06MI051	NP06MI051F	UO	5.569264277	*	4.109219643	***	no	Mother-Fetus	8	
NP06MI052	CK07MI035	OO	5.921581717	***	8.177587514	***	yes	Father-Son	8	7CN
NP06MI063	CK14MI023	OO	5.628770847	*	-2.579521145	ns	no	FP#		
NP06MI069	CS03MI013	OO	5.112659013	*	0.157192115	*	no	FP		
NP06MI077	NP06MI077F	OO	5.517589271	*	6.685760793	***	yes	Mother-Fetus	7CN	
NP07MI033	NP07MI033F	OO	5.318944549	*	5.041122174	***	no	Mother-Fetus	7CS	
NP07MI037	NP07MI037FA	OO	5.968106303	***	8.803691126	***	yes	Mother-Fetus###	7CS	
NP07MI037	NP07MI037FB	OO	7.337656291	***	9.44685602	***	yes	Mother-Fetus###	7CS	
NP07MI039	NP07MI039F	UO	7.233658428	***	9.124357001	***	yes	Mother-Fetus	7CS	
NP07MI040	01BC019	JJ	5.592389658	*	-2.033399991	(*)	no	FP#		
NP07MI042	CS08MI026	OU	6.685440103	***	8.713431863	***	yes	Father-Daughter	7CS	7CS
NP07MI051	NP07MI051F	OO	5.37463325	*	6.873540775	***	yes	Mother-Fetus	7CS	
NP07MI060	NP08MI011	OO	6.360368114	***	8.008819632	***	yes	Father-Son	7CS	9
NP07MI079	CK09MI011	OO	5.516177718	*	-1.137735308	(*)	no	FP		
NP08MI001	NP07MI048F	UO	6.389589647	***	10.12457409	***	yes	Father-Fetus	8	7CS
NP08MI011	NP11MI018	OO	6.354500944	***	7.486557406	***	yes	Father-Daughter	9	7CS
NP08MI022	CK09MI028	OU	5.481248436	*	8.549818499	***	yes	Father-Daughter	9	7CN
NP09MI011	CK15MI024	OO	7.172404249	***	10.7163976	***	yes	Son-Father	9	7CN
NP09MI013	CS08MI054	OO	5.036313933	(*)	7.783421948	***	yes	Father-Daughter	7WR	7CS
NP11MI040	09BC065	JJ	5.483783644	*	-0.539037746	*	no	FP#		
NP11MI044	05BC039	OO	6.164243401	***	8.430359054	***	yes	Mother-Son#.##	7CN	2

CS03MI039	15BC031	JJ	5.830596569	***	-6.471802267	ns	no	FP#		
CS03MI043	CS06MI039	OO	5.425884102	*	8.079097617	***	yes	Daughter-Father	7CS	7CS
CS05MI001	05BC071	OO	5.529056427	*	7.156328658	***	yes	Mother-Daughter##	7CS	7CS
CS05MI015	13BC080	OO	4.988290596	(*)	4.28114123	***	no	FP#		
CS05MI030	CS05MI050	OO	5.033179244	(*)	6.837755643	***	yes	Father-Son###	7CS	7CS
CS05MI030	06BC071	OO	5.691057503	***	7.603225165	***	yes	Father-Daughter###	7CS	7CS
CS06MI051	10BC055	JJ	5.361876013	*	-1.082572773	(*)	no	FP#		
CS06MI053	06BC075	JJ	5.514984404	*	8.624029545	***	yes	Mother-Daughter##	7CS	6
CS06MI053	14BC007	JJ	5.681854884	*	3.248027112	**	no	FP#		
CS07MI008	CS07MI008F	OO	4.989352893	(*)	7.817658405	***	yes	Mother-Fetus	7CS	
CS07MI037	CK05MI002	OO	6.104401408	***	7.892061975	***	yes	Daughter-Father	7CS	7CN
CS07MI052	CS07MI052F	OO	6.127618962	***	8.761445473	***	yes	Mother-Fetus	7CS	
CS07MI053	CS07MI053F	OO	7.617080593	***	11.16567979	***	yes	Mother-Fetus	7CS	
CS08MI030	CS08MI030F	OO	5.187757237	*	6.947101413	***	yes	Mother-Fetus	7CS	
CS08MI047	CS08MI047F	UO	6.320423521	***	9.702462631	***	yes	Mother-Fetus	7CS	
CS09MI019	CK05MI059	OO	8.064239622	***	7.035748071	***	no	FP#		
CS09MI029	CK08MI030	OO	5.510954467	*	6.383277712	***	yes	Daughter-Father	7CS	7CN
CS10MI007	03BC019	JJ	5.158823191	*	7.888686742	***	yes	Father-Daughter	7CS	6
CS10MI038	08BC128	JJ	5.508855758	*	2.129123897	**	no	FP#		
CS12MI029	CS13MI025	JJ	5.005715691	(*)	7.311716349	***	yes	Mother-Son##	7CS	7CS
CS12MI048	CK11MI059	OO	6.977507196	***	9.769846118	***	yes	Daughter-Father	7CS	7CN
CS13MI017	CK11MI005	JJ	5.950196149	***	-2.083597469	(*)	no	FP#		
CK04MI008	14BC073	OO	5.156473253	*	-8.639407134	ns	no	FP#		
CK05MI002	07BC085	OO	6.337566516	***	10.48917176	***	yes	Father-Daughter	7CN	11
CK05MI015	06BC003	JJ	5.300289974	*	-4.936920476	ns	no	FP		
CK05MI025	04BC059	OO	5.585343699	*	9.188445429	***	yes	Father-Daughter	7CN	7CS
CK07MI016	12BC027	JJ	5.373983062	*	3.38823831	**	no	FP#		
CK08MI017	12BC035	JJ	5.320138268	*	4.454807983	***	no	FP#		
CK08MI017	13BC035	JJ	5.32359816	*	3.736657547	***	no	FP		

CK08MI022	CK10MI006	JJ	5.399412647	*	8.560511909	***	yes	Mother-Son##	7CN	7CN
CK08MI031	CK10MI057	JJ	4.99864055	(*)	-1.70527731	(*)	no	FP		
CK09MI001	12BC082	JJ	5.150179042	*	-1.534344706	(*)	no	FP#		
CK09MI042	05BC043	OO	6.620542272	***	10.08875535	***	yes	Son-Father#	7CN	2
CK11MI053	09BC100	JJ	5.326476124	*	-0.714138344	*	no	FP#		
CK11MI054	CK15MI016	OO	5.60875643	*	2.780626215	**	no	FP		
CK11MI057	07BC086	JJ	5.505400506	*	-1.2705611	(*)	no	FP#		
CK12MI026	02BC096	JJ	5.380758881	*	-2.338407243	(*)	no	FP#		
CK13MI023	CK13MI023F	OO	6.183902815	***	9.48600141	***	yes	Mother-Fetus	7CN	
CK13MI029	09BC045	OO	6.899047818	***	10.27031846	***	yes	Father-Daughter	7CN	7CS
CK13MI053	CK13MI053F	OO	6.143216575	***	7.987463375	***	yes	Mother-Fetus	7CN	
CK14MI001	CK14MI001F	OO	5.010673554	(*)	8.455671291	***	yes	Mother-Fetus	7CN	
CK14MI013	14BC076	OO	5.128969356	*	7.193052062	***	yes	Mother-Daughter#.##	7CN	7CS
CK14MI041	12BC079	OO	4.972944433	(*)	6.966744977	***	yes	Son-Mother#.##	7CN	7CN
CK15MI002	CK15MI002F	OO	5.281425223	*	7.05822394	***	yes	Mother-Fetus	7CN	
CK15MI007	CK15MI007F	OO	6.197143623	***	9.179321399	***	yes	Mother-Fetus	7CN	
CK15MI023	06BC011	JJ	4.96592835	(*)	6.496122718	***	yes	FP#		
01BC041	04BC100	JJ	6.051870998	***	-4.261839023	ns	no	FP#		
02BC014	15BC011	JJ	4.921379986	(*)	3.708856137	***	no	FP#		
02BC019	09BC027	JU	4.935135272	(*)	2.570335446	**	no	FP#		
02BC099	15BC148	JJ	5,757088534	***	1,907859346	**	no	FP#		
03BC012	04BC001	JJ	4.946890553	(*)	-5.950002936	ns	no	FP		
03BC039	09BC046	JJ	5.238494504	*	-1.323994729	(*)	no	FP#		
03BC097	15BC157	JJ	5.77080971	***	6.06239283	***	no	Mother-Son##	6	6
03BC107	07BC059	JJ	5.607899925	*	1.641527957	**	no	FP#		
04BC008	14BC101	JJ	5.074361833	(*)	4.550839454	***	no	FP#		
06BC032	07BC059	JJ	6.320384276	***	5.719751379	***	no	Mother-Daughter##.###	6	6
06BC032	11BC049	JJ	5.97018077	***	6.152264307	***	no	Mother-Daughter##.###	6	6
06BC146	13BC083	JJ	5.069104285	(*)	6.105701592	***	yes	FP#		

07BC021	07BC111	JJ	5.253989451	*	1.413567849	**	no	FP		
07BC021	10BC123	JJ	6.312472248	***	6.94809719	***	yes	Mother-Son##	7CS	7CS
07BC059	11BC049	JJ	6.263725782	***	9.42126446	***	yes	Daughter-Mother##	6	6
07BC071	12BC068	JJ	5.626624701	*	4.319459213	***	no	FP		
07BC104	08BC019	JJ	4.979005062	(*)	6.581763676	***	yes	FP#		
07BC129	09BC074	JJ	5.022002254	(*)	-2.49142979	ns	no	FP#		
07BC142	08BC103	JJ	4.92493136	(*)	2.845874829	**	no	FP		
09BC024	09BC050	JJ	4.961277638	(*)	-1.885973737	(*)	no	FP		
10BC027	10BC109	JJ	5.690162171	***	1.124295205	*	no	FP		
11BC085	15BC049	JJ	4.990099531	(*)	3.020892717	**	no	FP#		
12BC116	15BC041	JJ	5.048218178	(*)	3.045667593	**	no	FP#		
14BC112	15BC059	JJ	5.619125037	*	1.105274626	*	no	FP		

Table 3: Comparison of estimated and observed values of False Discovery Rate (FDR) and Power (16 microsatellite loci; full data set n=4,554 n_f=53; only J stock n=1767 n_f=2; only O stock n=2366 n_f=49; J stock + unassigned n=2188 n_f=4; O stock + unassigned n=2787 n_f=51). FDR_O is calculated under the assumption that pairs' assignment as False Positives (FP) based on evaluation of 26 loci and biological data is correct. P_O is the percentage of Mother-Fetus pairs in the data set inferred to be P-O pairs by the Likelihood analysis.
P=power; E=estimated/expected; O=observed; n_f=number of fetusses in data set; #no mother-fetus pair detected
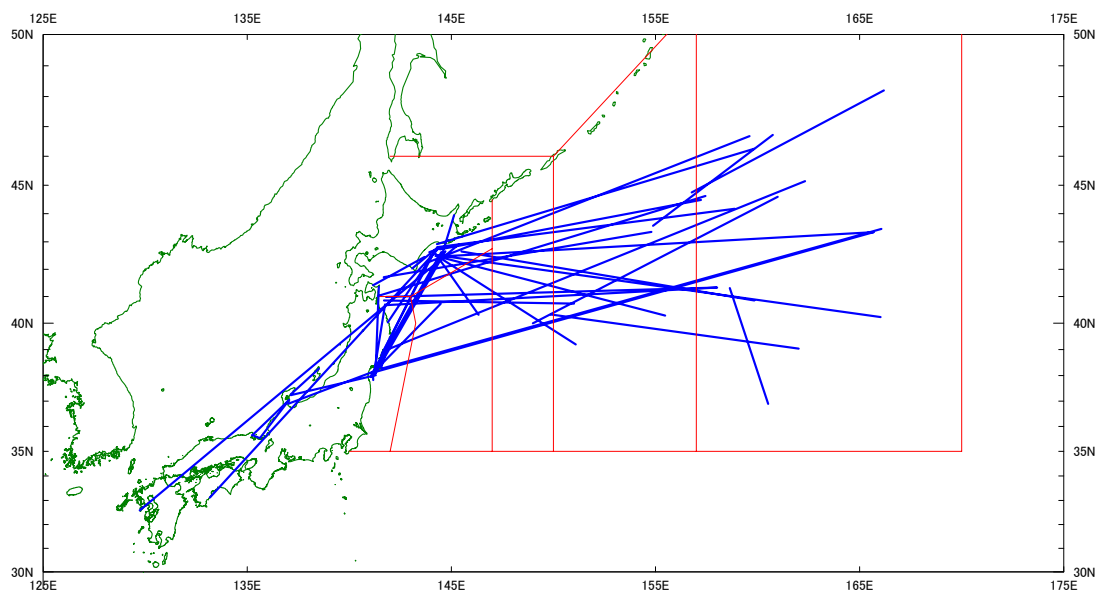
		Full dataset			Only J stock			Only O stock			J stock + unassigned			O stock + unassigned		
		FDR_O	P_E	P_O	FDR_O	P_E	P_O#	FDR_O	P_E	P_O	FDR_O	P_E	P_O#	FDR_O	P_E	P_O
FDR_E	0.1	0.468	0.437	0.500	0.000	0.117	0.000	0.258	0.555	0.347	0.200	0.177	0.000	0.236	0.530	0.373
FDR_E	0.05	0.383	0.383	0.444	0.000	0.102	0.000	0.133	0.413	0.306	0.000	0.146	0.000	0.133	0.395	0,333
FDR_E	0.01	0.179	0.240	0.277	0.000	0.102	0.000	0.103	0.272	0.163	0.000	0.146	0.000	0.065	0.264	0.176
FDR_E	0.001	0.179	0.240	0.277	0.000	0.102	0.000	0.103	0.272	0.163	0.000	0.146	0.000	0.065	0.264	0.176

Table 4: Observed Parent-Offspring pairs within and among stocks compared to expected values under a random distribution (significance evaluated with X^2 test)
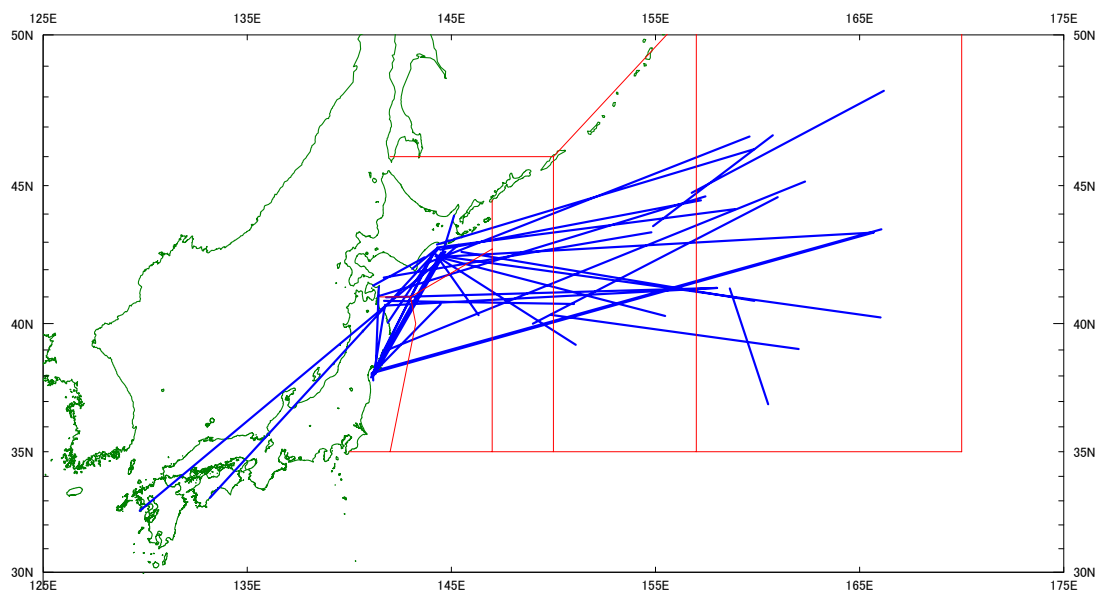
	Expected P-O pairs	Observed P-O pairs	p of X^2-Test	
Within J stock	11.29	9	0.460	ns
Within O stock	20.24	57	<0.001	***
Within unassigned	0.64	0	0.422	ns
Among J stock and O stock	30.24	0	<0.001	***
Among J stock + unassigned	5.38	0	0.016	*
Among O stock + unassigned	7.21	9	0.482	ns

Figure 1: Geographical representation of the validated 49 Parent-Offspring pairs (excluding Mother-Fetus pairs). The endpoints of any line represent the positions of the two individuals inferred to form a P-O pair.

A) All 49 pairs

(B) O stock pairs

(C) J stock pairs