PHYLOGEOGRAPHY OF NEPTUNE WHELK (*Neptunea arthritica*) SUGGESTS SEX-BIASED IMPACT OF TRIBUTYLTIN POLLUTION AND OVERFISHING AROUND NORTHERN JAPAN.

NORIKO AZUMA¹*, RICHARD M. MIRANDA², SEIJI GOSHIMA¹ AND SYUITI ABE³

¹Graduate School of Fisheries Sciences, Hokkaido University, 3-1-1, Minato-cho, Hakodate 041-8611, Japan
²Universidad Austral de Chile, Sede Puerto Montt, Región de Los Lagos, Chile.
³Sanriku Fisheries Research Center, Department of Revitalization for Sanriku-region, Iwate University, 3-75-1 Heita, Kamaishi 026-0001, Japan

Running head: PHYLOGEOGRAPHY OF NEPTUNEA WHELK

*Correspondence: Noriko Azuma, E-mail: nazu@fish.hokudai.ac.jp
The Neptune whelk, *Neptunea arthritica*, is a sublittoral sea snail from Pacific waters that has been a food resource and is commercially important for the coastal fisheries in northern Japan. This species showed a severe decline during the 1970s and 1980s, possibly because of overfishing, imposex caused by tributyltin (TBT) pollution and parasite infection. In the present study, we investigated genetic variation among the populations of *N. arthritica* from eight localities in northern Japan, including Hokkaido and Aomori, using a mitochondrial DNA (mtDNA) marker, partial sequence of the cytochrome c oxidase subunit I (COI) gene, to compare the obtained results with those from previous microsatellite analyses. We also addressed the evolutionary history of *N. arthritica* and human impact on the population genetic profiles of this species. The parsimony network showed 14 COI haplotypes separated into 2 groups (Groups A and B), with an intermediate haplotype connecting both of the groups. Among eight populations, six were fixed for only one or two haplotypes, and any geographic–genetic correlation was not found; they were probably affected by random drift of the mtDNA lineage. Thus, the results from mtDNA contrasted with those from previous microsatellite analysis, indicating that geographic structure was affected by the restricted gene flow between populations. Our results suggested that *N. arthritica* diversified into Groups A and B during the Pliocene; however, recent TBT pollution and size-selective fishing pressure have reduced genetic diversity and concealed the natural population structure. The present
study also suggested that human impact may cause long and possibly irreversible modification of ecosystems, particularly for species forming discrete and relatively small local populations, such as *N. arthritica*. Thus, the combined use of mtDNA and microsatellite genetic data provides a powerful tool to investigate the health of biodiversity in molluscs and shows that the output results of such analyses are of great interest for the conservation and management of molluscan species.
INTRODUCTION

*Neptunea arthritica* (Bernardi, 1857) is a dioecious gastropod with internal fertilisation and direct development in the sublittoral zone to a depth of a few tens of metres. The egg masses are deposited on hard substrata such as rocks and boulders, and it takes 3 years or more for maturation (Fujinaga, 2003). The typical *N. arthritica* (*N. arthritica arthritica*) is distributed in the Pacific Ocean, the Sea of Japan and the Sea of Okhotsk along the coasts of northern Japan and Sakhalin in southern Russia, whereas a subspecies *N. arthritica cumingii* (hereafter *N. a. cumingii*) is found from the western part of the Sea of Japan to the East China and Yellow Sea (Okutani, 2000), with the range partly overlapping with typical *N. arthritica*. Sea snails, including *N. artheritica*, have been a food resource and commercially important in the coastal fisheries in northern Japan (Mizushima & Torisawa, 2003); thus, several biological studies of *N. arthritica* have been mainly conducted for resource management (Kawai *et al.*, 1994; Fujinaga & Nakao, 1996; Suzuki *et al.*, 2002; Fujinaga, 2003; Fujinaga *et al.*, 2006; Miranda *et al.*, 2007 & 2009; Miranda, Fujinaga & Nakao, 2008; Lombardo & Goshima, 2010). However, none of the population genetic studies had appeared before our recent microsatellite DNA analysis in *N. arthritica* around Hokkaido (Azuma *et al.*, 2011). Using five loci of microsatellite DNA markers in seven populations of *N. arthritica* around Hokkaido, we suggested the restricted gene flow among populations with increasing genetic differentiation among populations separated by increasing geographic distances, i.e. following the
isolation-by-distance model (Azuma et al., 2011). The observed restricted gene flow between local populations was most probably influenced by the balance of the transport force of sea water and the low level of dispersal potential of this species (Azuma et al., 2011). Therefore, the suggested genetic structure was considered to be a result of natural distribution without strong anthropogenic disturbance.

However, the microsatellite data could not provide much knowledge regarding phylogeny and evolutionary history of this species in a palaeontological time scale. To reconstruct the evolutionary history of *N. arthritica*, we chose nucleotide sequence variation in the 5′ portion of the cytochrome *e* oxidase subunit I (COI) gene in the mitochondrial genome as a genetic marker. This marker was used in the present study for the following reasons: (1) Genetic markers from mtDNA have an advantage in genealogy analyses because they lack recombination and uniparental (maternal) inheritance (which results in the absence of heterozygotes); this makes it feasible to clarify lineages in comparison with markers from nuclear DNA (Harrison, 1989; Avise, 2000; Freeland, 2005). (2) The COI region examined in the present study showed sufficient variation within species and included a barcoding portion where sequence data were accumulated in many taxa; thus, it was used to compare sequences with those from other species. (3) mtDNA sequence data are available for molecular clock estimation, from which the divergence time of lineages can be estimated (Kumar, 2005).
Besides utility for the reconstruction of the evolutionary process, the mtDNA marker shows a higher ability to disclose the past bottleneck effect because of the small effective population size, which is a quarter of that of nuclear DNA (Moore, 1995). This indicates that mtDNA has sufficient sensitivity for detecting past population declines. Previous microsatellite DNA analyses for *N. arthritica* did not reveal any evidence of a recent decline in each population (Azuma *et al.*, 2011). However, around Hokkaido, *N. arthritica* showed a severe decline during the 1970s and 1980s, possibly because of overfishing, imposex caused by tributyltin (TBT) pollution and parasite infection (Kawai *et al.*, 1994; Fujinaga *et al.*, 2006; Miranda *et al.*, 2007 & 2009). Using the genetic profile of mtDNA, we can expect to detect such a recent decline better compared with when using microsatellites. In particular, if the main factors for population declines are severe in females, we can expect a drastic reduction in genetic diversity in mtDNA, which represents variability in the matriline. The skewed sex ratio caused by size-selective fishing has been reported in many fishery resource species (Rowe & Hutchings, 2003; Fenberg & Roy, 2008; Kendall & Quin, 2013). Similarly, size-selective harvesting, in which larger snails are caught, may cause more serious fishing pressure on females than on males in *N. arthritica* because the maturing size is larger in females than in males (Fujinaga, 2003; Miranda *et al.*, 2008). Imposed induced by TBT was observed in many species of gastropods modifying genitals and sterilizing females (Gibbs, 1996; Blackmore, 2000; Pavoni *et al.*, 2007; Bigatti *et al.*, 2009). Thus, human impact, overfishing,
and TBT pollution are likely to have affected populations in a sex-biased manner (more severe in females than in males), while parasite infection seemed to damage reproduction in both sexes (Miranda et al., 2009).

The present study aimed to genetically characterise *N. arthritica* populations around Hokkaido using an mtDNA marker and to compare the obtained results with those from previous microsatellite analyses to address the following two topics: (1) evolutionary history of *N. arthritica* and (2) human impact on population genetic profiles of this species. For (1), we tried to clarify the evolutionary scenario in the paleontological time scale using mtDNA data, which included genetic diversity, haplotype genealogy and spatial distribution of haplotypes. For (2), we analysed mtDNA data in comparison with results of microsatellite analyses. If the sign of genetic drift such as low level of genetic diversity within population appeared in mtDNA, definitely conflicting with the results from microsatellite DNA, we assumed sex-biased damage by human impact, which was more severe in females than in males in the examined populations.

**MATERIALS AND METHODS**

Specimens

We used 238 individuals of *N. arthritica* from seven locations in Hokkaido, namely Wakkanai (WA), Rumoi (RU), Kumaishi (KU) and Shiriuchi (SH) on the Sea of Japan coast; Toyoura (TO).
and Nemuro (NE) on the Pacific Ocean coast and Saroma (SA) on the Sea of Okhotsk coast, as well as from Aomori (AO) in northernmost Honshu (Table 1, Fig. 1). Hereafter, the term ‘sample’ is used for a group of individuals collected from each of the abovementioned localities, as representative of the local population. The samples were identical to those used for our previous microsatellite DNA analysis (Azuma et al., 2011), except for RU, which was recruited in the present study. Genomic DNA of the RU sample was extracted using the Pure Gene Kit (Qiagen) according to the manufacturer’s protocol, as described previously (Azuma et al., 2011), and used for analysis.

Nucleotide sequencing

The 5′ region of mtDNA COI was amplified by polymerase chain reaction (PCR) in a 30 μl reaction mixture containing template DNA (approximately 500 pg), dNTPs, a pair of primers [LCO1490 (5′-GGT CAA CAA ATC ATA AAG ATA TTG G-3′) and HCO2198 (5′-TAA TCA GGG TGA CCA AAA AAT CA-3′; Folmer et al., 1994)] and Taq DNA polymerase (Sigma), according to the manufacturer’s instructions. The thermal cycling profile included precycling denaturation at 95°C for 5 min, followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 45°C for 45 s and extension at 72°C for 45 s. After electrophoretic examination on a 2% agarose gel, the PCR products were purified with magnetic beads (AMPure, Agencourt), cycle-sequenced
using the abovementioned forward and reverse primers and the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and loaded onto an automated sequencer, ABI PRISM™ 3130 (Applied Biosystems). The obtained sequences of both directions were aligned and edited to 428 bp using DNASIS-Mac v.3.5 (Hitachi) and ClustalX 1.81 software (Thompson et al., 1997) for defining haplotypes and deposited in the DDBJ/Genbank database with accession Nos. AB432872–AB432884 and AB811355.

Molecular phylogeny

A phylogenetic tree of the obtained COI haplotypes was reconstructed using Bayesian algorithm in MrBayes 3.12 (Ronquist & Huelsenbeck, 2003). The sequences with high similarity (>92%) to the obtained data were searched by Basic Local Alignment Search Tool (BLAST) in the DDBJ/Genbank database. HQ834061 (in N. a. cumingii), FJ710085 (in N. arthritica) and FJ710084 (in N. a. cumingii) were found and included in the present phylogenetic tree and haplotype network analyses. A COI sequence from N. eulimata, accession No. EU883634, was used as an outgroup for phylogenetic analysis. We applied the substitution model GTR + G + I, which was recommended as the best fitting substitution model for our data set by jModelTest 0.1.1 (Posada, 2008; Guindon & Gascuel, 2003). In Bayesian analysis, the posterior probability distribution of trees was approximated by drawing a sample every 100 steps over 1,000,000 Markov chain Monte Carlo
(MCMC) cycles, in which the average standard deviation dropped to less than 0.00001, after
discarding a burn-in of 250,000 cycles. The length of burn-in was determined by the number of
cycles reaching the stability of log likelihood values. The haplotype genealogy within species was
resolved with a parsimony network using the TCS Network Program (Clement, Posada & Crandall,
2000) under 95% connection limit, with gaps as the 5th state.

The divergence time within species was estimated following the calibration detailed in
Nakano et al. (2010), which assumed that the subgenus *Barbitonia*, including *N. arthritica,*
diveredged from other *Neptunea* species approximately 11 million years ago (MYA) on the basis of
reliable fossil records of the oldest *Barbitonia.*

Population genetic analyses

We used a program package of Arlequin version 3.1 (Excoffier, Laval & Schneider, 2005) to
estimate haplotype (\( h \)) and nucleotide diversity (\( \pi \)) in each sample and to detect genetic
differentiation among samples by the calculation of pairwise \( F_{ST} \) (Weir & Cockerham, 1984).

Genetic differentiation between the samples was visualised on a two-dimensional surface by
non-metric multidimensional scaling (nMDS) plotting on the basis of pairwise \( F_{ST} \) using the
statistical software R version 2.9.0 (R Development Core Team). To test the significance of the
hierarchical population structure, analysis of molecular variance (AMOVA; Excoffier, Smouse &
Quattro, 1992) was conducted with Arlequin version 3.1 (Excoffier et al., 2005) assuming the three categories that were suggested by haplotype distribution and geography: 1. [WA, SA, RU] and [KU, SH, AO, TO, NE], 2. [WA, SA, RU, NE] and [KU, SH, AO, TO] and 3. [WA, SA, RU], [KU, SH, AO, TO] and [NE].

Evaluation of the isolation-by-distance (IBD) model (Wright, 1943) to assess the level of gene flow was performed using the abovementioned Arlequin program. For the IBD test, the geographic distance between sample locations was determined from the putative migration routes of whelks (Fig. 1). The distance matrix determined in this manner was compared with the \( F_{ST} \) matrix, and the significance of correlations was evaluated by the Mantel test.

RESULTS

**COI sequence variation and haplotype genealogy**

PCR amplification of approximately 650-bp fragments was not always successful in the examined whelk specimens, probably because of the low quality of extracted DNA. To eliminate unreliable sequences, a confirmed part of the 428-bp sequence was used for haplotype identification. Thus, 19 polymorphic nucleotide sites were found in the aligned sequences of COI from 238 analysed individuals, which defined a total of 14 haplotypes, *NACOIHI1–H13* and *NACOIAl* (Fig. 2). The
results of BLAST search revealed that the most frequent haplotype in our analysis, \textit{NACO1H1}, was identical to five 428-bp sequences in the DDBJ/Genbank database (accession Nos. JN053006, JN053006 and EU883627 from \textit{N. a. cumingii}; EU883629 from \textit{Neptunea} sp.1 and FJ710078 from \textit{N. arthritica}). In the BLAST search, we also found that the database sequence of FJ710085 for \textit{N. a. cumingii} was identical to the sequences of \textit{N. arthritica} (AB498776, AB498777 and AB498778) and \textit{N. a. cumingii} (FJ710083 and FJ710079). Thus, typical \textit{N. arthritica} and \textit{N. a. cumingii} shared at least two haplotypes, \textit{NACO1H1} and FJ710085 (Fig. 2). In the Bayesian tree (Fig. 2), all the haplotype sequences observed herein and those from databank, except for EU883634, were separated in three clusters, Group A (\textit{NACO1H1–5}, \textit{NACO1H7}, \textit{NACO1A1} and HQ83061 from databank), Group B (\textit{NACO1H6} and \textit{NACO1H8–13}) and the third group consisting of two other databank sequences, FJ710085 and FJ710084, whereas \textit{NACO1H10} was intermediate to these groups. Although the posterior probabilities for Group A and the third group (0.61 and 0.80, respectively) were not high enough to support the monophyly, the three groups were discriminated in a parsimony haplotype network (see below). The haplotype network (Fig. 3) was three forked, also showing two groups of haplotypes as seen in the Bayesian tree, Groups A and B, with core haplotypes (\textit{NACO1H1} and \textit{NACO1H6}) and derived haplotypes around core ones. \textit{NACO1H10} was present in the centre of the network, connecting Groups A and B and the third one containing databank sequences FJ710084 and
FJ710085. Several missing haplotypes appeared between groups, indicating lineage sorting within each group, and a star-like shape with core and derived haplotypes in each group suggested recent radiation.

Based on the GTR + G + I model, the genetic distance was estimated to be 6.7% between EU883634 (*Neptunea eulimata*) and *NACO1H1*. Given the 6.7% divergence for 11 MYA in the separation of *Barbitonia* from other *Neptunea*, the divergence rate per million years was estimated to be 0.609%. Considering the genetic distance of 1.6%–2.3% between Groups A and B, the divergence time between the two groups was estimated to be 4.67–2.65 MYA during the Pliocene.

The divergence time of haplotypes within each group (0.2%–0.4% difference from each other) was estimated to be approximately 0.3–0.65 MYA during the Pleistocene.

**Genetic population structure**

The haplotype distribution within samples is shown in Figure 4 and Supplementary Data. The haplotype *NACO1H1* was common among the examined samples, except for SA. The SA sample contained only *NACO1H5*, which was probably derived from *NACO1H1* with two substitutions (Fig. 3). Haplotypes from Group A occurred in every sample, whereas haplotypes from Group B were found in only two samples, KU and TO. Haplotype *NACO1H10*, connecting both the groups, occurred only in NE.
As shown in Table 1, the haplotype diversity ($h$) was moderate, and the nucleotide diversity ($\pi$) was low as a whole. Both $h$ and $\pi$ were the highest in KU, which had two Group A and six haplotypes from Group B (Supplementary Data), followed by WA, which had five haplotypes from Group A. The WA sample showed low $\pi$ because of a lack of haplotype from Group B. The third highest $h$ and the second highest $\pi$ were observed in TO. Both $h$ and $\pi$ were zero in three monomorphic samples, SA (only NACO1H5), RU (only NACO1H1) and SH (only NACO1H1).

This diversity profile was contrasting with the results of previous microsatellite marker analysis (Azuma et al., 2011; Table 1), in which the expected heterozygosity in each sample (0.577–0.729) was comparable with the total estimate (0.673). 

$F_{ST}$ analysis (Table 2) revealed that 20 out of 28 pairs of samples were genetically different (in bold letter); however, the difference/similarity pattern strikingly differed from the results of previous microsatellite analysis (Azuma et al., 2011). In nMDS plotting, the $F_{ST}$ estimates using microsatellite markers showed a correlation between the geographic and genetic structure (Fig. 5-A), whereas the genetic distance of mtDNA haplotypes between samples did not show a correlation with their spatial distribution (Fig. 5-B). The SH sample was distinctly separated from neighbouring KU and TO but completely overlapped with RU. On the other hand, KU and TO were in close proximity to each other. The SA sample was clearly distant from other populations, probably reflecting the exclusive occurrence of NACO1H5 but lack of NACO1H1, a major common
AMOVA failed to support any of the hierarchical structures in the category 1, 2 and 3 suggested by haplotype distribution and geography ($P = 0.15, 0.12$ and 0.51, respectively).

The Mantel test did not show a significant correlation between genetic ($F_{ST}$) and geographic distance ($P = 0.22$), indicating that the examined *N. arthritica* populations did not follow the IBD model with the current mtDNA data.

**DISCUSSION**

*mtDNA phylogeny and phylogeography of N. arthritica*

Haplotypic distribution was heterogeneous among the localities, and the localisation of lineages was probably due to the historical dispersal pattern. Considering the limited distribution of Group B, only in KU and TO in southern Hokkaido, and the results of $F_{ST}$ analysis using microsatellite DNA (Fig. 5), the genetic differentiation between southern and northern population is plausible. The haplotype *NACO1H10*, present in the centre of the haplotype network and thus potentially ancestral to all other haplotypes observed herein, was found only in NE, and it may indicate that the species possibly originated from the east, the Kuril Islands. However, we could not delineate a certain structure or evolutionary process because $F_{ST}$ analyses, AMOVA and IBD test failed to capture a reasonable geographic–genetic structure in mtDNA. The loss of genetic diversity
in some populations may hide the structure in these analyses. It is likely that SH may have possessed haplotypes from Group B in the past, similar to neighbouring KU and TO. If haplotypes from Group B had remained in SH populations, the geographic structure would have been easily described as north–south differentiation. The possible cause of the genetic loss in SH, genetic drift, is discussed later, with comparison of results from the present mtDNA and previous microsatellite DNA analyses.

Sharing of some Group A sequences and those retrieved from databank (FJ710084 and FJ710085) in both *N. arthritica arthritica* and *N. a. cumingii* indicates that the latter is genetically indistinguishable from the former. Sometimes, the name of *N. cumingii* appeared as a full species (WoRMS Editorial Board, 2014); however, Hou *et al.* (2013) suggested that *N. cumingii* and *N. a. cumingii* are the same species in molecular phylogenetic analysis using mtDNA and nuclear DNA. Combined results of Hou *et al.* (2013) and the present study suggest that *N. a. cumingi* and so-called *N. cumingii* are not full species but subspecies or a geographic form of *N. arthritica*.

Discordance of mtDNA and microsatellite DNA phylogeography of *N. arthritica*

The population genetics inferred from mtDNA analysis was not consistent with that inferred from previous microsatellite DNA analysis, and recent genetic drift is the most plausible reason for the discordance. The appropriate sample collection in the present study was proved by various
alleles and HWE in each sample in microsatellite analyses using same individuals; thus, the discordance of the results from two markers was not due to the artefact in the field or laboratory works but reflected the actual property of *N. arthritica* around Hokkaido.

In our previous microsatellite DNA analyses, each of the examined sample (local population) of *N. arthritica* showed genetic diversity (*H*<sub>E</sub> = 0.577 to 0.729) that was comparable with the total diversity estimation (*H*<sub>E</sub> = 0.673), and the population structure was correlated to geography (Supplementary File of the present study; figs. 3 and 4 in Azuma et al., 2011). In contrast, the present mtDNA analyses provided different population genetic profiles: three of eight samples (SA, RU and SH) were monomorphic, showing extremely lower diversity (*h* = 0 and *π* = 0) than the total estimation (*h* = 0.57 and *π* = 0.0061) (Table 1). Such situation can be generally considered to be a result of the bottleneck by a founder effect or genetic drift in a small size of the local population. In some species of a low dispersal ability and small local population size, the local population is likely fixed for one or a few haplotypes, as seen in the Japanese crayfish (Koizumi et al., 2012). Such species usually showed apparent genetic–geographic correlation, and it seems reasonable that the low dispersal ability caused both low genetic diversity within the population and geographic structure among populations, probably by stepwise migration in their evolutionary history. However, in *N. arthritica*, the departure from IBD (Mantel test), negative AMOVAs and unreasonable nMDS plotting pattern based on *F*<sub>ST</sub> revealed no genetic–geographic correlation. In
summary, it is conceivable that the observed mtDNA phylogeographical pattern in *N. arthritica* was influenced by very recent genetic drift. Genetic drift stochastically left a small number of genotypes (Harrison, 1989), and the natural genetic structure related to geography may be hidden after the drift. Thus, genetic drift could be a reason for the genetic–geographic inconsistency as well as for the lack of genetic diversity in some *N. arthritica* populations. However, if the bottleneck occurred a long time ago, genetic diversity should be more or less recovered even in mtDNA by gene flow, as suggested by our microsatellites analysis (Azuma et al., 2011). Thus the genetic drift was thought to be recent. Possible causes of the genetic drift in *N. arthritica* include natural biotic and abiotic factors, e.g. predation, parasitism, disease, change in climate and topology and human impact such as exploiting, environmental modification and pollution. Among these, the human impact, overfishing and imposex caused by TBT pollution, was considered to be the most plausible cause of the contrasting results obtained from mtDNA and microsatellite analyses. In the *N. arthritica* population around Hokkaido, overfishing and TBT pollution were reported to be specific causes of the extreme population decline in the 1970s and 1980s (Fujinaga *et al.*, 2006; Miranda *et al.*, 2007 &2009), and they were reported to be surely related to the skewed sex ratio in the reproductive stage. Recently, Toews & Brelsford (2012) reviewed 126 studies exhibiting discordant biogeography of mtDNA and nuclear DNA (mito–nuclear discordance) in animal species. They concluded that the most frequent reason for mito–nuclear discordance was
sex-biased asymmetries, including sex-biased offspring production, and that very rare cases were able to be solely explained by genetic drift in both sexes and the small effective population size in mtDNA. The sex-biased asymmetry could be the reason for the striking mito–nuclear discordance in *N. arthritica*. The sex ratio (male/female) in prosobranch gastropods has been generally reported to be 1:1 (Hughes, 1986; Power & Keegan, 2001; Ilano, Fujinaga & Nakao, 2003); the ratio in *N. arthritica* was reported to be 0.82 in 2003–2004 in Lake Saroma (Miranda *et al.*, 2009). This suggests that the female number is surely equal to or a little greater than the male number in stable *N. arthritica* populations. Miranda *et al.* (2009) also observed that almost all normal adult females (i.e. without imposex or parasites) had abundant sperm in their capsule gland in April–June 2003 and June 2004 in Lake Saroma, suggesting that all mature females join the annual reproduction. Thus, it is not likely that a lesser number of females than males produce offspring under normal condition. However, if imposex occurs in females, it causes sex-biased asymmetry in reproduction, a decrease in the number of females involved in reproduction. Thus we conclude that the recent imposex caused by TBT pollution and severe matrilineal decline was considered to be the most plausible cause of mito–nuclear discordance in *N. arthritica*. Fujinaga *et al.* (2006) reported virtual recovery from imposex in *N. arthritica* populations around Hokkaido after banning TBT use. In addition, previous microsatellite DNA analyses in *N. arthritica* (Azuma *et al.*, 2011) and *Nucalla lapillus* (Colson & Hughes, 2004) revealed a substantial level of genetic diversity in each
population of the two species, suggesting a rapid recovery of genetic diversity in the nuclear genome from the genetic disturbance of TBT-induced imposex. Nevertheless, the present mtDNA analysis suggested that a bottleneck effect caused by TBT pollution is still responsible for the lack of diversity in matrilines of *N. arthritica* around Hokkaido.

Another cause of deficiency of mtDNA genetic diversity is overfishing. Fujinaga *et al.* (2006) described an exceptionally high sex ratio (male/female > 1.4) in the four localities around Hokkaido in 2002, attributing it to fishing pressure. If the fishing pressure is higher on females than on males, it reduces the effective population size of females to a greater extent than that of males.

For example, in Hiyama district in southern Hokkaido, Fisherman’s Cooperative Association prohibits the catch of sea snails, mainly *N. arthritica*, of a small size, i.e. less than 6 cm of shell height. The mature size of shell height was reported to be 50 mm in males and 60 mm in females at Usu Cove (Fujinaga, 2003) and 60 mm in males and 75 mm in females in Lake Saroma (Miranda *et al.*, 2009). This may suggest that the fishery restriction as that seen in Hiyama area caused sex-selective fishery, in which more reproductive females would be caught than males. Other factors, anthropogenic transplantation and/or population decline by parasite infection, are not likely to be responsible for the striking mito–nuclear discordance in *N. arthritica* because these factors should affect microsatellite DNA variation as well as mtDNA. The level of genetic impact of TBT pollution and/or overfishing probably differs between localities. Some of the examined populations,
WA, KU and TO, have maintained a high level of genetic diversity with regard to the haplotype diversity index $h$, and it may indicate that the negative impact was low in these populations.

The SA population in Lake Saroma showed a low level of genetic diversity both in mtDNA and microsatellite analyses as the monomorphic haplotype component in mtDNA and the lowest genetic diversity in microsatellites. Lack of genetic diversity in both markers is attributable to specific reasons in this population, namely a founder effect in recent population establishment and parasite infection, in addition to TBT pollution and overfishing. In SA, haplotype $NACOHI$, which was ubiquitous and the most abundant in total samples, was not found, and all individuals had haplotype $NACOIH5$, which was found in only two individuals in WA, the closest to SA among the populations examined in the present study (Fig. 1). The founder effect by recent establishment of this population probably caused this particular phenomenon. Lake Saroma is connected with the Sea of Okhotsk by a channel, which was first opened in 1929. The diatom assemblages and data of sedimentary ages from bore hole samples revealed that the salinity of Lake Saroma had increased in 1929 (Kashima, 1996), thereby indicating that the time of migration or introduction of $N. arthritica$, the species unable to survive in the low salinity, was after 1929. It is likely that many or all of the founders derived from the source population had $NACOIH5$ at that time. Because two individuals in WA also had this haplotype, it is not likely that $NACOIH5$ originally evolved in SA. Of course, imposex by TBT pollution, overfishing and parasite infection threatened this population.
as well as other populations, and it might enhance the founder effect, reducing genetic diversity.

Severe parasite infection was observed in SA (Miranda, 2009), and it may be more severe in SA than in other habitat of *N. arthritica* around Hokkaido because of the enclosed water, a feature different from other coastal habitats.

*Evolutionary history of N. arthritica*

As mentioned above, the loss of genetic diversity in many populations makes it difficult to reconstruct the evolutionary history of *N. arthritica* around Hokkaido. Thus, the following is a very rough sketch of the evolution of this species. The main diversification of species into Groups A and B was estimated during the Pliocene, 4.67–2.65 MYA, at the onset of global cooling, and this dating does not contradict the fossil record of *N. arthritica* in a deposit of the Pliocene (Amano, 1997). By the late Pliocene, endemic speciation of many molluscan species, which characterise the Omma–Manganji fauna (Otuka, 1939; Amano, 2007), occurred in the Sea of Japan. This event was influenced by the environmental change in the Sea of Japan, which was semi-closed by a land bridge connecting the Korea Peninsula and Kyushu and lifting backbone range of mountains on the Japanese Archipelago (Chinzei, 1978). The diversification of *N. arthritica* may be enhanced by such environmental change. The eurythermal capacity of *N. arthritica* may have allowed it to survive through the drastic climate and topological changes in the Pleistocene, as hypothesised by Amano.
(1997), while many sympatric Neptunea species went extinct. The several missing haplotypes in each branch of Groups A and B in the haplotype network (Fig. 3) may suggest that *N. arthritica* had suffered climate oscillation as well as other species and lost many lineages. The present haplotype distribution, which showed that Group B was found only in southern Hokkaido, may suggest that Groups A and B were allopatrically established and contacted later. The diversification within each group started at the middle of Pleistocene, as inferred by genetic diversity between core haplotypes and derived ones (0.2%); however, the star-like shape of each group in the haplotype network may indicate more recent radiation. In the present study, there appeared to be no reproductive isolation between Groups A and B within each population because no deviation from the Hardy–Weinberg equilibrium was found at any of the microsatellite DNA loci examined in KU or TO (Appendix in Azuma *et al.*, 2011), both of which included Groups A and B (Fig. 4 and Supplementary Data). The restricted gene flow found in the microsatellite DNA analyses suggested that the settled local populations were somewhat isolated from each other and a small number of migrants may be responsible for gene flow. Each of local population has evolved on the balance of such isolation and migration; in other words, local decline was rescued by recruitment from neighbouring populations. Some local populations have shrunk since the 1970s (Kawai *et al.*, 1994; Fujinaga *et al.*, 2006) because of TBT pollution and/or overfishing. This reduction in the population size is reflected by poor genetic diversity in mtDNA and it has erased important genetic evidence to detect
precise evolutionary history in this species. At present, the population size appears to be recovering in each locality; however, the genetic diversity that once decreased in the matriline is likely unrecovered. The matrilineal diversity in each population may partly recover with gene flow in the future; however, genetic recovery definitely needs much longer time than the recovery of the population size.

Conclusion

In the present study, comparison of mtDNA data with microsatellite DNA indicated that sex-biased asymmetry in population genetics was probably affected by anthropogenic pollution and fishing pressure in *N. arthritica*. The legislation prohibiting TBT usage as an antifouling agent for coastal boats and aquaculture constructions was implemented in 1990 in Japan, and many countries, including Japan, ratified a total TBT ban proposed by International Marine Organization. However, the effect may persist for a considerable period of time. Both TBT pollution and overfishing were stopped around Hokkaido more than 15 years before sample collection for the present study; however, loss of genetic diversity was not recovered in the matriline of *N. arthritica*. It is important to know that human impact may cause long and possibly irreversible modification in ecosystems, particularly in species forming discrete and relatively small local populations, such as *N. arthritica*. 
ACKNOWLEDGEMENTS

This work was partly supported by MEXT Revitalization Project for the Creation of Fisheries Research and Education Center in Sanriku. The authors would like to thank Enago (www.enago.jp) for the English language review.
REFERENCES


Figure 1. Map of sampling locations of *Neptunea arthritica* in northern Japan, Wakkanai (WA), Rumoi (RU), Kumaishi (KU), Shiriuchi (SH), Toyoura (TO), Aomori (AO), Nemuro (NE) and Saroma (SA). Dashes indicate putative migration pathways.

Figure 2. Fifty-percent majority-rule Bayesian tree inferred from partial mtDNA CO1 sequences of *Neptunea arthritica* using the GTR + G + I model. Bold and italic OTU indicates the haplotype found in the present study, and the others are the accession numbers of sequences retrieved from DDBJ/GenBank. The tree was rooted using EU883634 from *N. eulimata* as an outgroup. Nodal numbers represent Bayesian posterior probability values.

Figure 3. Parsimony network of the mtDNA COI haplotypes of *Neptunea arthritica*. Open circles indicate a haplotype observed in the present study, and the circle size reflects haplotype abundance (number of individuals that had the haplotype). Squares and closed circles indicate a sequence retrieved from the database and a missing haplotype, respectively. A solid line between circle/square indicates a single nucleotide substitution.

Figure 4. Distribution of the mtDNA COI haplotypes in each sampling locality of *Neptunea*
arthritica. Note that six of eight samples have only one or two haplotypes, and distant RU and SH share only NACO1H1.

Figure 5. The non-metric multidimensional scaling (nMDS) plotting of Neptunea arthritica samples with pairwise $F_{ST}$ values. A: based on five loci of microsatellite DNA markers (Azuma et al., 2011), B: based on a 428-bp sequence of partial mtDNA COI. In A, the horizontal long scattering plot, which is consistent with geographic relationships between samples, suggests a population structure with a one-dimensional genetic cline, from eastern and northeastern Hokkaido to southern Hokkaido and northernmost Honshu (Azuma et al., 2011).
Table 1. Informations of Neptune whelk samples analyzed in the present study.
Sample name, Collection date, Sample size (number of individuals), and diversity indices, haplotype diversity ($h$) and nucleotide diversity ($\pi$) estimated in partial COI sequence, and mean expected heterozygosity ($H_e$) estimated in five loci of microsatellite (Azuma et al., 2011). Sample of RU were not analyzed with microsatellites because of poor amplification in PCR.

<table>
<thead>
<tr>
<th>Sample name</th>
<th>Collection date (year, month)</th>
<th>Sample size</th>
<th>haplotype diversity ($h$)</th>
<th>Nucleotide diversity ($\pi$)</th>
<th>Expected heterozygosity ($H_e$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NE</td>
<td>2007, 03</td>
<td>30</td>
<td>0.33±0.08</td>
<td>0.0030±0.0013</td>
<td>0.643</td>
</tr>
<tr>
<td>SA</td>
<td>2006, 09</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0.577</td>
</tr>
<tr>
<td>WA</td>
<td>2007, 09</td>
<td>30</td>
<td>0.59±0.08</td>
<td>0.0019±0.0015</td>
<td>0.726</td>
</tr>
<tr>
<td>RU</td>
<td>2007, 10</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>KU</td>
<td>2006, 11</td>
<td>30</td>
<td>0.71±0.05</td>
<td>0.0088±0.0028</td>
<td>0.717</td>
</tr>
<tr>
<td>SH</td>
<td>2006, 09</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0.720</td>
</tr>
<tr>
<td>TO</td>
<td>2007, 03</td>
<td>30</td>
<td>0.48±0.05</td>
<td>0.0078±0.0019</td>
<td>0.729</td>
</tr>
<tr>
<td>AO</td>
<td>2007, 11</td>
<td>28</td>
<td>0.13±0.08</td>
<td>0.0032±0.0006</td>
<td>0.606</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>238</td>
<td>0.57±0.03</td>
<td>0.0061±0.0036</td>
<td>0.673</td>
</tr>
</tbody>
</table>
**Table 2.** Pairwise $F_{ST}$ between *Neptuna arthritica* samples based on partial COI sequence. Bold letter indicates significant deviation from 0 at $p<0.01$ after Bonferroni correction.

<table>
<thead>
<tr>
<th></th>
<th>NE</th>
<th>SA</th>
<th>WA</th>
<th>RU</th>
<th>KU</th>
<th>SH</th>
<th>TO</th>
</tr>
</thead>
<tbody>
<tr>
<td>NE</td>
<td></td>
<td>0.763</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SA</td>
<td>0.153</td>
<td></td>
<td>0.813</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WA</td>
<td>0.172</td>
<td>1.000</td>
<td></td>
<td>0.118</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RU</td>
<td>0.422</td>
<td>0.714</td>
<td>0.549</td>
<td>0.590</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KU</td>
<td>0.172</td>
<td>1.000</td>
<td>0.118</td>
<td>0.000</td>
<td>0.590</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH</td>
<td>0.445</td>
<td>0.739</td>
<td>0.574</td>
<td>0.621</td>
<td>-0.0315</td>
<td>0.621</td>
<td></td>
</tr>
<tr>
<td>TO</td>
<td>0.156</td>
<td>0.968</td>
<td>0.103</td>
<td>0.041</td>
<td>0.574</td>
<td>0.041</td>
<td>0.604</td>
</tr>
</tbody>
</table>
Haplotype frequency of partial COI (number of individuals which showed the haplotype) in each sample. Bold letter indicates the highest frequency in each sample.

<table>
<thead>
<tr>
<th></th>
<th>NACO1</th>
<th>NACO1</th>
<th>NACO1</th>
<th>NACO1</th>
<th>NACO1</th>
<th>NACO1</th>
<th>NACO1</th>
<th>NACO1</th>
<th>NACO1H</th>
<th>NACO1H</th>
<th>NACO1H</th>
<th>NACO1H</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>NE</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>SA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>WA</td>
<td>18</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>RU</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>KU</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>SH</td>
<td>30</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>TO</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>AO</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>Total</td>
<td>149</td>
<td>2</td>
<td>7</td>
<td>1</td>
<td>32</td>
<td>32</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>238</td>
</tr>
</tbody>
</table>
* found in only *N. a. cumingii*
**shared by *N. arthritica* and *N. a. cumingii*
Group A

* found in only *N. a. cumingii*

**shared by *N. arthritica* and *N. a. cumingii***