HOKKAIDO UNIVERSITY

Title	Population genetics and dynamics of the giant Pacific octopus (Enteroctopus dofleini) in the North Pacific
Author(s)	岸本, 真琴
Degree Grantor	北海道大学
Degree Name	博士(水産科学)
Dissertation Number	甲第16562号
Issue Date	2025-09-25
DOI	https://doi.org/10.14943/doctoral.k16562
Doc URL	https://hdl.handle.net/2115/98172
Type	doctoral thesis
File Information	Makoto_Kishimoto.pdf, 全文

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Population genetics and dynamics of the giant Pacific octopus (*Enteroctopus dofleini*) in the North Pacific

（北太平洋におけるミズダコ[*Enteroctopus dofleini*]の集団遺伝学と個体群動態に関する研究）

Division of Marine Bioresource and Environmental Science
Graduate School of Fisheries Sciences
Hokkaido University
北海道大学大学院水産科学院
海洋生物資源科学専攻

KISHIMOTO Makoto
岸本真琴

2025年

Table of Contents

Table of Contents ii

List of Tables iv

List of Figures v

Acknowledgements vi

Chapter 1: General introduction 1

Chapter 2: Genetic diversity and population structure of giant Pacific octopus (*Enteroctopus dofleini*) based on two mitochondrial DNA genes 8

2.1 Introduction 8

2.2 Materials and Methods 9

2.2.1 Sample collection 9

2.2.2 DNA extraction, amplification, and sequencing 11

2.2.3 Data analysis 12

2.3 Results 15

2.3.1 Cytochrome c oxidase I (COI) 15

2.3.2 Cytochrome b (CYTB) 25

2.4 Discussion 27

Chapter 3: Population genetic analysis of giant Pacific octopus (*Enteroctopus dofleini*) around Japan and estimation of recent genetic exchange based on RAD-sequencing and Microsatellite analyses 32

3.1 Introduction 32

3.2 Materials and Methods 34

3.2.1 Samples 34

3.2.2 ddRAD-seq library construction 35

3.2.3 Single nucleotide polymorphism (SNP) detection 36

3.2.4 Individual-based analyses 37
3.2.5 Population-based analyses 39
3.2.6 Migration analyses 41
3.2.7 Microsatellite analyses 42
3.3 Results 43
3.3.1 Population structure 43
3.3.2 Genetic diversity and population analyses 45
3.3.3 Estimated migration 55
3.4 Discussion 55
Chapter 4: Population dynamics and phylogenetic analysis of giant Pacific octopus (*Enteroctopus dofleini*) in the Northeast Pacific and Northwest Pacific 61
4.1 Introduction 61
4.2 Material and Methods 64
4.2.1 Demographic analysis 64
4.2.2 Estimation of divergence times 65
4.3 Results 68
4.3.1 Population demography 68
4.3.1.1 Northwest Pacific 68
4.3.1.2 Northeast Pacific 70
4.3.2 Phylogenetic analysis 70
4.4 Discussion 73
Chapter 5: General discussion 78
References 84
Appendices 100

List of Tables

Table 1 Sample size, haplotype number and diversity, nucleotide diversity, Tajima's *D* and Fu's F_S for *Enteroctopus dofleini* from sampling sites in Japan and pooled data from the Northeast Pacific 17

Table 2 Pairwise Φ_{ST} scores and p-scores between sampling sites for the cytochrome oxidase I (COI) gene 21

Table 3 Pairwise Φ_{ST} scores and p-scores between sampling sites for the cytochrome *b* (CYTB) gene 22

Table 4 Results of analysis of molecular variance in *Enteroctopus dofleini* for the Northwest Pacific and Northeast Pacific groups 24

Table 5 Sampling sites and genetic diversity of *Enteroctopus dofleini* based on ddRAD-seq. 48

Table 6 Sampling sites and genetic diversity of *Enteroctopus dofleini* based on the seven microsatellite markers 49

Table 7 Pairwise F_{ST} scores for all population pairs excluding Asamushi and Fukui 51

Table 8 Estimated migration rates between each sampling site by BA3-SNPs for Dataset POP 56

Table 9 Summary of demographic indices for the Northwest Pacific (NWP) and Northeast Pacific (NEP) 69

List of Figures

Figure 1 Life cycle of the giant Pacific octopus (*Enteroctopus dofleini*) near Japan..... 3

Figure 2 Sampling sites of giant Pacific octopus (*Enteroctopus dofleini*) in Japan and of reference samples 10

Figure 3 Median-joining networks for mtDNA haplotypes of giant Pacific octopus (*Enteroctopus dofleini*) from the Northeast Pacific and Northwest Pacific for the cytochrome oxidase I gene and cytochrome b gene 16

Figure 4 Maximum likelihood trees for mtDNA sequence data from giant Pacific octopus (*Enteroctopus dofleini*) for the cytochrome oxidase I gene and cytochrome b gene..... 19

Figure 5 Nonmetric multidimensional scaling with Φ_{ST} scores from the Northwest Pacific and Northweast Pacific for each gene (COI and CYTB). 23

Figure 6 Principal component analysis (PCA) of population structure using Dataset IND. 44

Figure 7 Neighbor-Net network using Dataset IND..... 46

Figure 8 Results of STRUCTURE analysis 47

Figure 9 Relationship between genetic distance (F_{ST} / (1 – F_{ST})) and geographic distance from Mantel tests 52

Figure 10 Principal coordinates analysis (PCoA) of Dataset POP..... 53

Figure 11 Principal coordinates analysis (PCoA) of microsatellite data..... 54

Figure 12 Demographic analyses of the NWP and NEP populations 71

Figure 13 Phylogenetic tree of 14cephalopods based on Bayesian inference of 13 protein coding regions from the mitochondrial gene with *Todarodes pacificus* as the outgroup. 72

Figure 14 Schematic representation of proposed changes in historical distribution of *Enteroctopus dofleini* based on the results of this study 77

Acknowledgements

First, I would like to express my heartfelt gratitude to my academic advisor, John R. Bower, for his invaluable support and guidance throughout my studies, and I deeply apologize for any concern I may have caused. I am truly grateful for the freedom I was given to conduct my research.

I extend my sincere appreciation to the teachers who advised me and reviewed my research: Prof. Hideaki Kudo, Prof. Takafumi Fujimoto, and Asst. Prof. Shingo Akita. I am also deeply thankful to Asst. Prof. Noriko Azuma, Tokyo University of Agriculture, and Prof. Hitoshi Suzuki for their valuable support and feedback on this research. Additionally, I wish to express my gratitude to Prof. Hiroyuki Munehara and Asst. Prof. Shingo Akita for kindly allowing the use of laboratory facilities for analysis.

I thank the following organizations for facilitating the collection of tissue samples: Fisheries Research and Education Agency, Hokkaido Research Organization Fisheries Research Department, Echizen Matsushima Aquarium, and Asamushi Aquarium. I also thank the fishers who granted permission to collect samples.

Grants from the Research Institute of Marine Invertebrates (Grant No. KO2022-10) and Hokkaido University DX Doctoral Fellowship Program (JST SPRING, Grant Number JPMJSP2119) made this research possible.

I am also incredibly thankful to the current and former members of the Laboratory of Marine Bioresource and Environmental Science and to Ms. Kiko Shibata for their valuable advice on this research.

Finally, a heartfelt thank you to my family for their warm and unwavering support throughout my research journey.

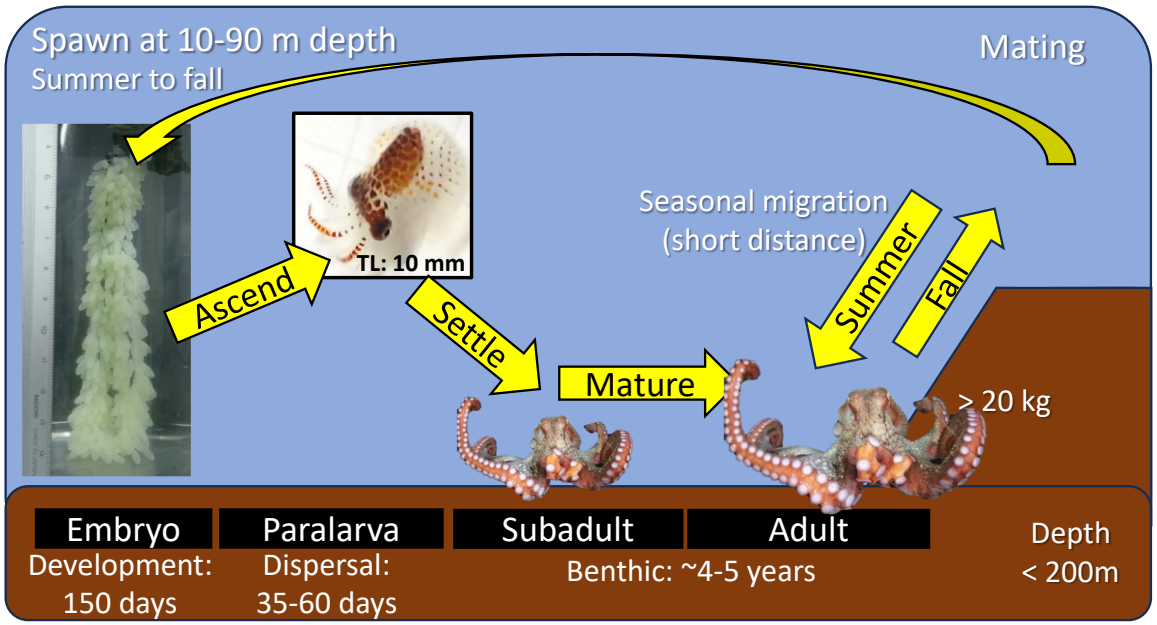
Chapter 1: General introduction

In marine organisms, the distribution of genetic diversity is shaped by both contemporary and historical processes. Gene flow is important in shaping within-species genetic variation patterns and is mediated through dispersal (Bohonak 1999). Larval dispersal plays a key role in structuring populations. Species with large numbers of small, widely dispersing, pelagic larvae often exhibit genetically homogeneous populations (Palumbi 1992, 1994). Extended pelagic larval durations allow wider dispersal, making it less likely for clear genetic population structures to form (*e.g.*, hair crab [*Erimacrus isenbeckii*], Azuma et al. 2008; sentinel crab [*Macrophthalmus japonicus*], Kobayashi et al. 2023). Dispersal is also affected by other factors such as currents and behavior. For example, Roura et al. (2019) showed that in coastal cephalopods in the North Atlantic, dispersal pathways and distances are affected by the depth distribution and size of the paralarvae. This indicates that various factors influence the genetic structure, so the relationship between dispersal and genetic population structure can be complex (Cowen & Sponaugle 2009; Esser et al. 2023).

Historical processes also influence phylogeographic patterns. During glacial periods in Earth's history, ice sheets expanded across vast areas of both polar regions and

sea levels dropped, which affected the survival of marine species (Tanaka & Takahashi 2005; Mann & Gaglioti 2024). In the North Pacific, some species, such as the Pacific cod (*Gadus macrocephalus*), show distinct genetic structures between the Northwest Pacific (NWP) and Northeast Pacific (NEP), which are thought to be caused by a separation of the species' range into distinct populations by the decline or southward shift of northern populations during the last glacial maximum (LGM) (Canino et al. 2010). For example, the file dogwinkle (*Nucella lima*) shows higher genetic diversity in the NWP than in the NEP based on demographic change and fossil record, which suggest a climate-related extinction in the NEP (Cox et al. 2014). Such patterns vary widely depending on the species' ecology and the location of refugia (*e.g.*, lined shore crab [*Pachygrapsus crassipes*], Cassone & Boulding 2006; red king crab [*Paralithodes camtschaticus*], Grant & Cheng 2012; Sitka periwinkle [*Littorina sitkana*], Azuma & Chiba 2017; Pacific herring [*Clupea pallasii*], Liu et al. 2012; sablefish [*Anoplopoma fimbria*], Orozco-Ruiz et al. 2023).

In cephalopods, particularly octopods, the dispersal ability varies depending on the size and morphology of the eggs and paralarvae (Boletzky 2003; Villanueva et al. 2016). One species with dispersive planktonic paralarvae is the giant Pacific octopus (*Enteroctopus dofleini*), whose hatchlings measure 9.5–10.1 mm in total length and settle on the bottom about 35–60 days after hatching (Okubo 1980, 1989; Fig. 1).



Cf. Okubo 1980, 1989, Kubodera 1991, Boletzky 2003, Noro & Sakurai 2012, Villanueva et al. 2016

Figure 1. Life cycle of the giant Pacific octopus (*Enteroctopus dofleini*) near Japan. TL, total length.

E. dofleini is the world's largest octopus and the most abundant species in coastal waters of the North Pacific. It occurs widely along the Pacific Rim from Mexico, to the Gulf of Alaska, Aleutian Islands, Bering Sea, Sea of Okhotsk, and Sea of Japan (Norman et al. 2016; Bower et al. 2024). In Japan, it commonly undertakes seasonal migrations between intertidal and deeper (>200 m) waters, but these movements are usually over short distances, and there is no evidence of population-wide migrations (*e.g.*, Kanamaru & Yamashita 1969; Sato & Yorita 1999; Noro & Sakurai 2012; Joh et al. 2017). Sano (2010) used geographical fishery data and the similarity in temporal changes in catch amounts around Hokkaido, Japan, during 1985–2004 to identify its main habitats. He partitioned the fishing grounds into 11 areas, suggesting that these areas reflect the seabed characteristics preferred by settled octopuses, such as rocky reefs, and inferred that steep slopes and sandy areas likely serve as habitat boundaries. These 11 areas are currently used to monitor and evaluate the fishery in Hokkaido (Wakkanai Fisheries Research Institute 2022).

Molecular approaches are increasingly being used to investigate the life histories and population dynamics of marine species (Avise 2000; Cowen & Sponaugle 2009; Hellberg 2009). Genetic information is important for fisheries management and conservation, and information about the population structure can provide clues about gene flow, migration, and demography, such as recent geographic separation. Yet, despite the

commercial importance of *E. dofleini* in Japan, little is known about its genetic diversity or population dynamics. Studies of its genetic population structure have been limited to the Gulf of Alaska (Toussaint et al. 2012a; Barry et al. 2013). Data on large-scale geographic patterns of genetic variation are also limited, but an unpublished study during 2005–06 of the population structure of *E. dofleini* in the North Pacific based on analyses of the mitochondrial cytochrome c oxidase subunit I (COI) gene and some microsatellite markers by P. R. Rigby and colleagues found evidence suggesting the occurrence of four populations including two in Russia and one each in Japan and the United States (Y. Shirayama, personal communication, 28 October 2024).

Cryptic species are groups of two or more taxa that were previously classified as a single nominal species due to their nearly identical appearance, which makes them difficult to distinguish morphologically (Hending 2025). They have been reported in many squids (*e.g.*, Sales et al. 2017; Fernández-Álvarez et al. 2021; Morgan et al. 2024). In octopuses, the *Octopus vulgaris* species complex is thought to comprise about three cryptic species (*O. vulgaris sensu stricto*, *O. sinensis*, and *O. vulgaris* type II) (Warnke et al. 2004; Leite et al. 2008; Amor et al. 2017). Toussaint et al. (2012a) reported the presence of a cryptic species of *E. dofleini* in Prince William Sound, Alaska, based on four nuclear and mitochondrial markers, and it was later confirmed using 11 microsatellite markers and

morphological characters (Hollenbeck & Scheel 2017; Hollenbeck et al. 2017). In Japan, Misawa et al. (2021) also reported this cryptic species from Tohoku (n = 4) and COI sequences (655 bp) matched those from the NEP. However, with only four individuals reported in Japan, the distribution of this cryptic species in Japan is unclear. Understanding the distribution of cryptic species is crucial for accurate conservation and management.

Octopuses are a critical component of marine ecosystems. They prey on a wide variety of organisms, including crabs, shellfish, and fish, while they themselves are preyed upon by higher predators such as marine mammals and sharks, positioning them in the middle of the ecological pyramid (Sauer et al. 2019). They are also important commercially in fisheries. The demand for octopuses is increasing worldwide, but the amount of catch has continued to decline (FAO FishStat 2025). To help meet this increasing demand, in recent years, studies have been conducted on *O. vulgaris* and *O. sinensis* to identify prey species for paralarvae and develop efficient rearing methods in closed systems to achieve full-life cycle aquaculture (*e.g.*, Vaz-Pires et al. 2004; Vidal 2014; Dan et al. 2018). Genetic knowledge is not only useful for understanding and conserving wild populations, but it may also aid in selecting desirable traits and optimizing breeding conditions when aquaculture is implemented in the future.

In this study, I used genetic approaches to better understand the population structure and phylogeny of *E. dofleini* in the North Pacific. In Chapter 2, I examined the genetic diversity and population structure of *E. dofleini* in northern Japan and evaluated the genetic differences of this species between the NWP and the NEP using two mitochondrial DNA markers: COI and cytochrome b (CYTB). In Chapter 3, to investigate population relationships in more detail using a large number of single-nucleotide polymorphisms (SNPs) and more diverged regions from nuclear regions, restriction site-associated DNA sequencing (RAD-seq) and microsatellite analyses were employed for population genetic analysis. Furthermore, in Chapter 4, the demographic history and phylogenetic relationships of populations in the NWP and NEP were reconstructed to explore the historical change on this region. The results of this study will provide fundamental and important insights for better management of *E. dofleini* fisheries and understanding of future fluctuations in resource abundance.

Chapter 2: Genetic diversity and population structure of giant Pacific octopus (*Enteroctopus dofleini*) based on two mitochondrial DNA genes

2.1 Introduction

Enteroctopus dofleini is broadly distributed across the North Pacific, but genetic studies of this species have focused primarily on phylogenetics, with samples from the NEP and NWP regions often treated as a single group (Takumiya et al. 2005; Kaneko et al. 2011; Ibáñez et al. 2016; Lima et al. 2020b). In a study of the population structure of *E. dofleini* in the NEP, Barry et al. (2013) showed that reference samples from the NWP formed a separate clade, but this result was not discussed. Due to the lack of studies published on the population genetics in *E. dofleini* in the NWP, the population structure in this region is uncertain.

Chapter 2 describes the results of my study to clarify the genetic diversity and population structure of *E. dofleini* in northern Japan and evaluate genetic differences of this species between the NWP (Japan and South Korea) and the NEP (Alaska and Canada) using two mitochondrial DNA markers, COI and CYTB, to compare with previous studies and detect cryptic species. These markers were chosen because COI is a widely used and

important barcode region in DNA barcoding and CYTB was shown in a previous study to have high diversity (Toussaint et al. 2012a).

2.2 Materials and Methods

2.2.1 Sample collection

A total of 261 specimens of *E. dofleini* were collected from Hokkaido and Honshu, Japan (Fig. 2). They included 237 adults and subadults collected in June–August 2022 at six sites: Hidaka, Iburi, Mombetsu, Obira, Shiriya, and Soya. The sites were selected based on Sano (2010) to include different fishing grounds currently used to monitor catches. The other 24 specimens comprised adults and subadults from Mori (June 2020–May 2022) and paralarvae and eggs from Asamushi (December 2020–June 2021) and Fukui (June 2020–October 2021). Specimens were collected by fishers, and precise geographic (latitude and longitude) data of the collection sites were not recorded.

From each adult and subadult specimen, a small piece of muscle tissue was cut from the arm tips or wall of the esophagus. Paralarvae and eggs were preserved whole. All samples were preserved in 70% ethanol at room temperature until DNA extraction.

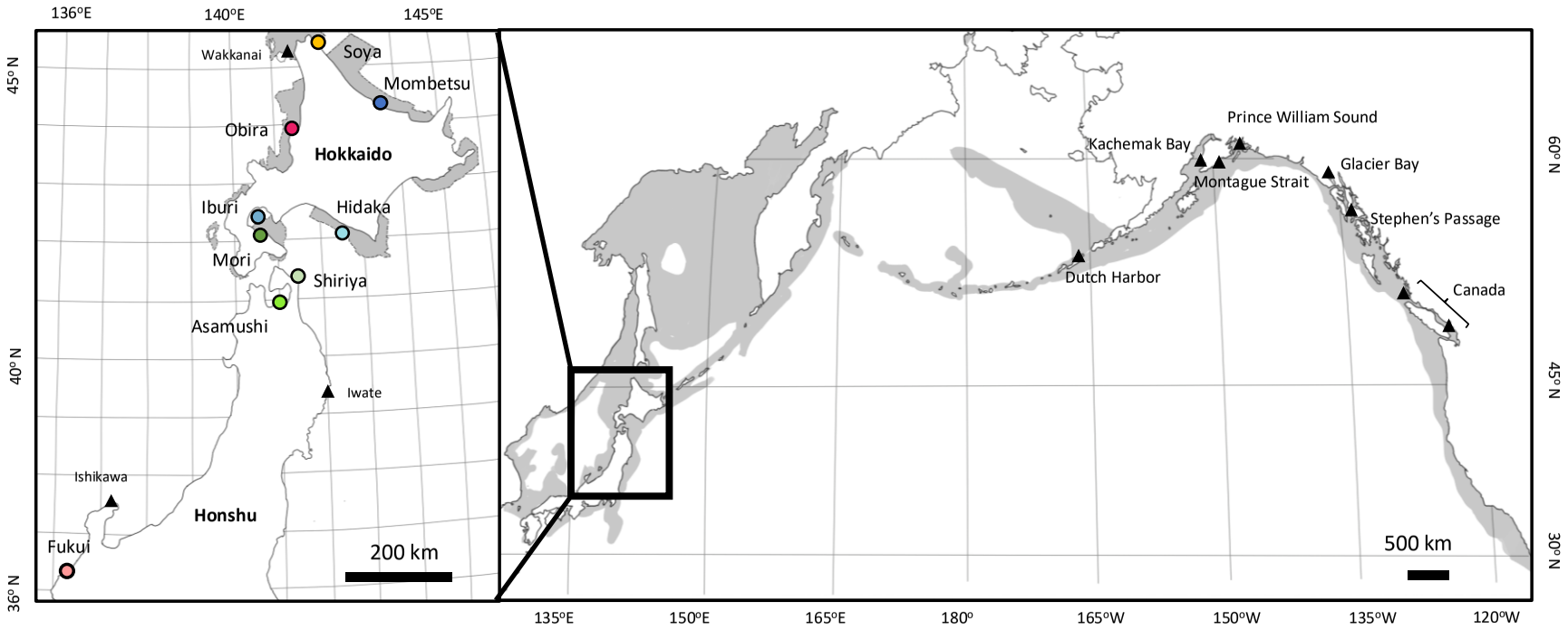


Figure 2. Sampling sites of giant Pacific octopus (*Enteroctopus dofleini*) in Japan (colored circles) and of reference samples (triangles). Left panel shows Northwest Pacific (NWP) collection and reference sample sites. Grey areas in left panel show the 11 areas now used to monitor catches (Sano 2010). Right panel shows Northeast Pacific (NEP) stations. Triangles indicate locations of reference samples. Locations of sampling sites in South Korea are not known, so triangles are not shown. Grey area in right panel shows the distribution of *E. dofleini* (adapted from Bower et al. 2024).

2.2.2 DNA extraction, amplification, and sequencing

Total genomic DNA was extracted using the NucleoSpin DNA RapidLyse kit (Takara Bio, Kusatsu, Japan) following the manufacturer's protocol. The quality of extracted DNA was tested using the GeneQuant Pro RNA/DNA Calculator (ver. 4082 2.6) (Amersham Pharmacia Biotech, Buckinghamshire, England) to confirm that the concentration and quality were suitable for polymerase chain reaction (PCR).

Partial sequences of the COI and CYTB regions were each amplified using two primer pairs: Forward-LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and Reverse-HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') for COI (Folmer et al. 1994), and Forward (5'-GGAACGAAGAATAGCATAAGC-3') and Reverse (5'-ATATGCTTAGTTCATTACG-3') for CYTB (Toussaint et al. 2012a). PCR amplification was conducted in a reaction of 10 μL reaction mixture comprising 1 μL of template DNA (10–100 ng/μL), 5 μL of Emerald Taq DNA polymerase (Takara Bio, Kusatsu, Japan), and 1 μL of each of the forward and reverse primers (10 pmol/μL each). The following thermal conditions were used: initial denaturation at 94℃ for 2 min followed by 35 cycles of denaturation at 94℃ for 45 s, annealing at 50℃ for 45 s and extension at 72℃ for 90 s, followed by a final extension at 72℃ for 30 min. The PCR products were visualized using agarose gel electrophoresis, purified using the NucleoSpin Gel and PCR Clean-up kit

(Takara Bio, Kusatsu, Japan) following the manufacturer's instructions, and sent to Macrogen Japan (Tokyo, Japan) for sequencing.

2.2.3 Data analysis

DNA sequences were aligned and edited using MEGA software (ver. 7.0) (Kumar et al. 2016). Reference sequences downloaded from GenBank were included in the statistical analyses. Reference sequences for COI were from Japan (n = 3; 1 each from Wakkanai in Hokkaido, Ishikawa and Iwate in Honshu), Alaska (n = 136; 49 in Kachemak Bay, KB, 46 in Dutch Harbor, DH, 33 in Glacier Bay, GB, 5 in Prince William Sound, PWS, 2 in Stephen's Passage, SP, and 1 in Montague Strait, MS), South Korea (n = 3), and Canada (n = 2); reference sequences for CYTB were from Alaska (n = 116; KB = 58, PWS = 32, DH = 20, and MS = 6) and South Korea (n = 3). Because of the smaller number of samples from Canada and South Korea, they were treated as a single group. Information about reference sequences is shown in Appendix A and B. Sequences of haplotypes newly nominated in the present study were deposited in DDBJ (Accession no.: LC807055–LC807068, LC807721–LC807734).

Molecular diversity indices of each sampling site and total of them including number of haplotypes, haplotype diversity (h), and nucleotide diversity (π) were calculated

using DnaSP software (ver. 6.12.03) (Rozas et al. 2017). The indices were not calculated for Fukui or Asamushi because of the small sample sizes. The diversity indices were also calculated for the NEP by pooling previously published sequence data from Alaska and Canada. Similarly, I pooled sample data from Japan and Korea as the NWP. Before the analyses, sequences were trimmed to 517 bp for COI and 597 bp in CYTB. Haplotypes that were assumed to be a lineage of a cryptic species (Toussaint et al. 2012a; Barry et al. 2013) were excluded from diversity analyses of both genes.

Since the correspondence between individual samples in the NEP reference sequences for COI and CYTB was unclear, subsequent analyses were conducted separately for each region. To infer the evolutionary relationships between haplotypes, a median-joining network was constructed using PopArt software (ver. 1.7) (Leigh & Bryant 2015). Molecular phylogenetic trees were constructed using the Maximum Likelihood (ML) method and Bayesian (BI) analyses. The best fit models for nucleotide substitution of each codon position were selected based on the Akaike Information Criterion (AIC) for ML using MEGA and Bayesian information criterion (BIC) for the BI analyses with Kakusan4 software (Tanabe 2007). The ML tree was constructed using RAxML ver. 8.2.9 (Stamatakis 2014) with 1,000 bootstrap replicates using the following models: for COI, the TN93 model for the first and second codon and the General Time Reversible (GTR)+G

model for the third codon; and for CYTB, the TN93 model for the first codon, the Hasegawa-Kishino-Yano (HKY)+G model for the second codon, and HKY model for the third codon. BI analysis was performed using MrBayes ver. 3.2.2. (Ronquist et al. 2012) under the HKY+G model for COI and CYTB. In the BI analysis, four independent Markov chain Monte Carlo (MCMC) chains were run for 2,000,000 generations, sampling every 1,000th generation. The first 25% of trees were removed as burn-in. The convergence within chains was assessed using Tracer ver. 1.6 (Rambaut et al. 2014) to determine if parameter values reached effective samples sizes (ESS) > 200. These phylogenetic trees were constructed using one sequence selected from each haplotype. Parts of the complete mitochondrial sequences from species in a closely related genus, *Muusoctopus johnsonianus* (Accession no. ON367815) and *M. leioderma* (Accession no. PP760091), were used as the outgroup.

To examine large-scale regional differences in the North Pacific, pairwise Φ_{ST} values, Tajima's *D* values (Tajima 1989) and Fu's F_S values (Fu 1997) were calculated for both regional groups (NWP and NEP). The statistical significance of all measures was verified using a permutation test in ARLEQUIN ver. 3.5.2.2 (Excoffier & Lischer 2010). To compare the likelihood based on Φ_{ST} scores between collection sites, nonmetric multidimensional scaling (nMDS) was performed with R (R Core Team 2024) using the

"vegan" package (Oksanen et al. 2025). To avoid negative value, the Φ_{ST} scores were converted to Slatkin linearized Φ_{ST} (Slatkin 1995). Hierarchical analysis of molecular variance (AMOVA) was conducted using ARLEQUIN ver. 3.5.2.2 to assess genetic divergence within and among the NEP and NEP.

2.3 Results

2.3.1 Cytochrome c oxidase I (COI)

A 604 bp segment of the COI region was successfully sequenced in 260 of the 261 individuals collected in this study (Table 1). After each sequence was trimmed to 517 bp, 11 haplotypes with 11 polymorphic sites were detected, including 4 parsimony-informative sites. The haplotype network with reference sequences showed the haplotypes were closely related (Fig. 3A). Haplotypes EdCO_1 and EdCO_2 were the most abundant around Japan, shared in 170 and 80 individuals, respectively, and occurred at all sites except Asamushi and Fukui. Haplotype EdCO_9 was common in Alaska but rare in Japan, appearing in one sample from Hidaka. Haplotype EdCO_12, which was connected to EdCO_1 and EdCO_2 with one and two substitutions, respectively, was also shared in the NEP (Canada) and NWP (Wakkanai, Hokkaido). Other less common haplotypes in the NWP were connected to EdCO_1 or EdCO_2 with one or two substitutions.

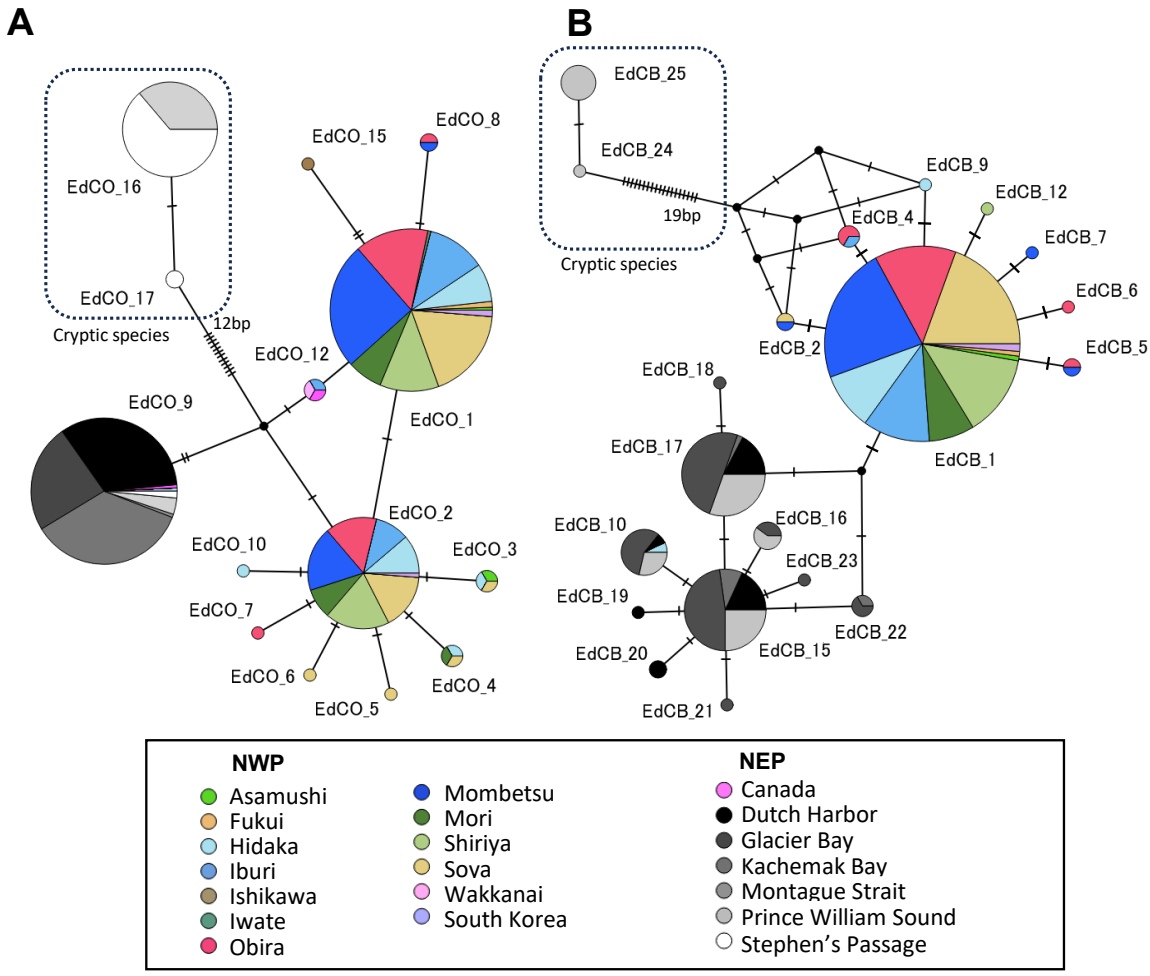


Figure 3. Median-joining networks for mtDNA haplotypes of giant Pacific octopus (*Enteroctopus dofleini*) from the Northeast Pacific (NEP) and Northwest Pacific (NWP) for the cytochrome oxidase I (COI) gene (A) and cytochrome *b* (CYTB) gene (B). Circle size is proportional to the number of individuals with each haplotype. Circle colors indicate areas in which the haplotypes were present and correspond to the colors in Figure 1. Note that one individual from Hidaka had specific haplotypes (EdCO_9 in COI and EdCB_10 in CYTB) that were otherwise found only in the NEP. EdCO_16 and 17, and EdCB_24 and 25 were a cryptic species (shown as enclosed by a dashed line) based on Toussaint 2012a and Barry 2013.

Table 1. Sample size (number of examined individuals), haplotype number and diversity (*h*), nucleotide diversity (π), Tajima's *D* and Fu's F_S for *Enteroctopus dofleini* from each sampling sites and pooled data from the Northwest Pacific (NWP) and Northeast Pacific (NEP). Population genetic analysis was not conducted for Fukui and Asamushi, where sample numbers were small.

Site	Sample size	Haplotype number	Haplotype diversity ($h \pm$ SD)	Nucleotide diversity ($\pi \pm$ SD)	Tajima's *D*	Fu's F_S
COI						
Mombetsu	59	3	0.411 ± 0.058	0.00081 ± 0.00012	–0.043	0.107
Soya	48	6	0.519 ± 0.064	0.00123 ± 0.00020	–1.069	*–2.508
Obira	39	4	0.506 ± 0.062	0.00108 ± 0.00017	–0.466	–0.753
Shiriya	35	2	0.504 ± 0.031	0.00098 ± 0.00006	1.598	1.749
Hidaka	26	6	0.649 ± 0.067	0.00190 ± 0.00045	–1.423	–1.905
Iburi	29	3	0.463 ± 0.080	0.00093 ± 0.00018	–0.108	–0.028
Mori	20	3	0.542 ± 0.076	0.00117 ± 0.00022	0.173	0.153
Fukui	2	1				
Asamushi	2	2				
This study	260	11	0.497 ± 0.024	0.00112 ± 0.00008	*–1.570	**–7.071
NWP	266	12	0.503 ± 0.024	0.00115 ± 0.00015	**–1.723	**–7.157
NEP	138	2	0.015 ± 0.014	0.00008 ± 0.00080	**–1.562	–1.316
CYTB						
Mombetsu	60	4	0.098 ± 0.053	0.00017 ± 0.00009	*–1.676	**–4.657
Soya	48	2	0.042 ± 0.040	0.00007 ± 0.00007	–1.107	*–1.602
Obira	38	4	0.201 ± 0.085	0.00035 ± 0.00015	*–1.566	**–3.100
Shiriya	35	2	0.057 ± 0.053	0.00010 ± 0.00009	*–1.136	–1.339
Hidaka	26	3	0.151 ± 0.093	0.00064 ± 0.00048	**–2.002	–0.527
Iburi	29	2	0.069 ± 0.063	0.00012 ± 0.00011	–1.149	–1.184
Mori	19	1				
Fukui	2	1				
Asamushi	2	1				
This study	259	9	0.091 ± 0.025	0.00019 ±0.00007	**–2.168	**–15.015
NWP	262	9	0.090 ± 0.024	0.00020 ± 0.00010	**–2.166	**–15.082
NEP	116	10	0.691 ± 0.026	0.00150 ± 0.00010	–1.126	*–4.264

* $p < 0.05$, ** $p < 0.01$. Significant values are highlighted in bold font.

For the COI region, 12 haplotypes were detected in the NWP and 2 in the NEP except for cryptic species. In the NWP, the overall haplotype diversity (h ± standard deviation [SD]) was 0.503 ± 0.024, ranging at each site from 0.411 to 0.649, and the overall nucleotide diversity (π ± SD) was 0.00115 ± 0.00015, ranging at each site from 0.00081 to 0.00190 (n = 266) (Table 1). Both scores were highest in Hidaka (h ± SD = 0.649 ± 0.067, π ± SD = 0.00190 ± 0.00045, n = 26). In the previously published sequence data from the NEP, only two haplotypes were detected (h ± SD = 0.015 ± 0.014, π ± SD = 0.00008 ± 0.00080, n =138). The haplotype frequency at each sampling site is shown in Appendix A.

Analysis using the ML and BI methods yielded a phylogenetic tree with two branches (Fig. 4A). The bootstrap value for the cryptic species was high which supported the independence of this clade. However, the low bootstrap values showed the phylogenetic relationships with the 11 haplotypes from the NWP and the NEP were unclear.

All pairwise comparisons between the NEP and NWP showed high pairwise Φ_{ST} scores (range: 0.391–0.992), and in 30 out of the 40 pairwise comparisons showed significant differences among most NWP sampling sites, but this result is considered

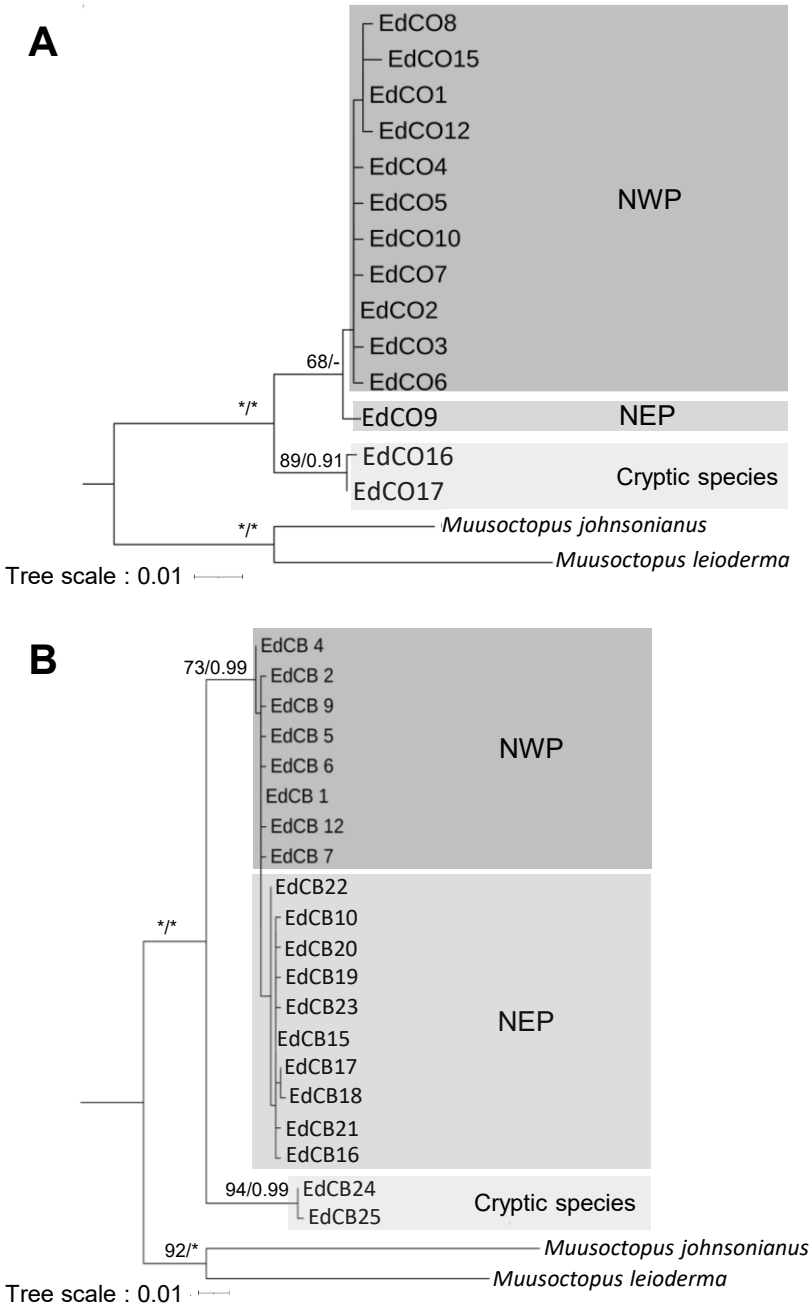


Figure 4. Maximum likelihood trees for mtDNA sequence data from giant Pacific octopus (*Enteroctopus dofleini*) for the cytochrome oxidase I (COI) gene (A) and cytochrome *b* (CYTB) gene (B). Trees include the lineage of cryptic species (EdCO16 and 17 in COI, EdCB24 and 25 in CYTB) and the outgroup *Octopus vulgaris*. Numbers near branches are bootstrap values and posterior probabilities. Only bootstrap values > 65 and posterior probabilities > 0.9 are shown. The bootstrap values = 100 and posterior probabilities = 1 are shown as *. Scale bar = number of nucleotide substitutions per nucleotide site.

inconclusive because of the small sample size (Table 2). For the three samples from South Korea, genetic differences with samples from the NWP were not significant and were not considered attributable to the small sample size. The South Korean samples shared haplotypes with the Japanese populations (EdCO_1 and EdCO_2 for COI). Pairwise comparisons within each group (NWP and NEP) showed no significant differentiations.

In the nMDS analysis, no groups showing correlation with specific regions were observed within the NWP population (Fig. 5). Mombetsu and Hidaka were positioned at the most distant ends of the plot, however, this arrangement did not directly reflect geographic distances. When compared with the NEP population, the NEP and NWP were separated into two distinct groups along nMDS1. The hierarchical population structure comprising the NWP and NEP groups were supported by the AMOVA analysis results (Table 4). The high variance component among groups (89.06%, $p < 0.01$) and low component of among sites within groups (0.66%, $p = 0.005$) and within sites (10.27%, $p < 0.01$) supported the regional division into the NWP and NEP and the low differentiation among sites in each regional group.

Tajima's *D* values were negative except for 1.598 in Shiriya and 0.173 in Mori, but did not significantly differ from zero at any site (range: – 1.423 to 1.598) (Table 1). The overall Tajima's *D* value in NWP was negative (–1.723) and significantly different from

Table 2. Pairwise Φ_{ST} scores (above diagonal) and p scores (below diagonal) between sampling sites for the cytochrome oxidase I (COI) gene.

		Northwest Pacific (NWP)								Northeast Pacific (NEP)				
CO1		Soya	Obira	Mombetsu	Hidaka	Iburi	Mori	Shiriya	South Korea	Canada	DH	GB	KB	PWS
NWP	Soya		−0.018	0.004	0.008	−0.012	−0.029	−0.012	−0.210	0.663	**0.913**	**0.900**	**0.916**	**0.852**
	Obira	0.966		−0.007	0.020	−0.019	−0.024	−0.008	−0.215	0.687	**0.931**	**0.919**	**0.933**	**0.870**
	Mombetsu	0.230	0.456		0.083	−0.022	0.018	0.043	−0.190	0.757	**0.938**	**0.929**	**0.939**	**0.900**
	Hidaka	0.251	0.143	0.013		0.042	−0.023	−0.014	−0.173	0.503	**0.896**	**0.875**	**0.900**	**0.776**
	Iburi	0.584	0.745	0.829	0.082		−0.006	0.018	−0.212	0.700	**0.949**	**0.939**	**0.951**	**0.890**
	Mori	0.934	0.720	0.193	0.747	0.471		−0.034	−0.220	0.644	**0.949**	**0.937**	**0.951**	**0.868**
	Shiriya	0.618	0.483	0.103	0.438	0.231	0.859		−0.195	0.702	**0.939**	**0.927**	**0.941**	**0.879**
	South Korea	1.000	1.000	1.000	1.000	1.000	1.000	1.000		0.391	0.992	**0.989**	**0.992**	0.938
NEP	Canada	0.002	0.001	0.001	0.012	0.005	0.011	0.001	0.170		0.026	0.027	0.123	0.314
	DH	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	0.000	0.918		0.000	1.000	1.000
	GB	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	0.889	1.000		1.000	1.000
	KB	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	0.923	0.000	0.000		1.000
	PWS	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	0.018	0.474	0.000	0.000	0.000	

Bold font indicates a significant deviation from zero at $p < 0.01$ level after Bonferroni correction. Abbreviations: DH, Dutch Harbor; GB, Glacier Bay; KB, Kachemak Bay; PWS, Prince William Sound.

Table 3. Pairwise Φ_{ST} scores (above diagonal) and p scores (below diagonal) between sampling sites for cytochrome *b* (CYTB) gene.

		Northwest Pacific (NWP)								Northeast Pacific (NEP)			
CYTB		Soya	Obira	Mombetsu	Hidaka	Iburi	Mori	Shiriya	South Korea	DH	KB	MS	PWS
NWP	Soya		0.016	−0.011	0.019	0.004	−0.023	0.001	−0.199	**0.959**	**0.809**	**0.892**	**0.858**
	Obira	0.096		0.008	0.009	−0.019	−0.010	0.009	−0.189	**0.900**	**0.777**	**0.843**	**0.815**
	Mombetsu	1.000	0.246		0.017	−0.003	−0.024	−0.003	−0.200	**0.943**	**0.815**	**0.890**	**0.860**
	Hidaka	0.196	0.248	0.222		0.003	−0.013	0.008	−0.198	**0.833**	**0.727**	**0.780**	**0.758**
	Iburi	0.612	1.000	0.797	0.334		−0.015	0.001	−0.198	**0.937**	**0.776**	**0.854**	**0.820**
	Mori	1.000	0.771	1.000	1.000	1.000		−0.019	0.000	**0.942**	**0.759**	**0.837**	**0.801**
	Shiriya	0.667	0.446	0.836	0.275	0.692	1.000		−0.199	0.946	**0.787**	**0.869**	**0.834**
	South Korea	1.000	1.000	1.000	1.000	1.000	1.000	1.000		0.827	0.699	0.733	0.719
NEP	DH	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	0.010	0.827		0.011	0.359	0.043
	KB	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	0.699	0.357		0.584	−0.018
	MS	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	0.733	0.006	−0.011		−0.009
	PWS	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	**> 0.00001**	0.719	0.214	0.899	0.460	

Bold font indicates a significant deviation from zero at $p < 0.01$ level after Bonferroni correction. Abbreviations: DH, Dutch Harbor; KB, Kachemak Bay; MS, Montague Strait; PWS, Prince William Sound.

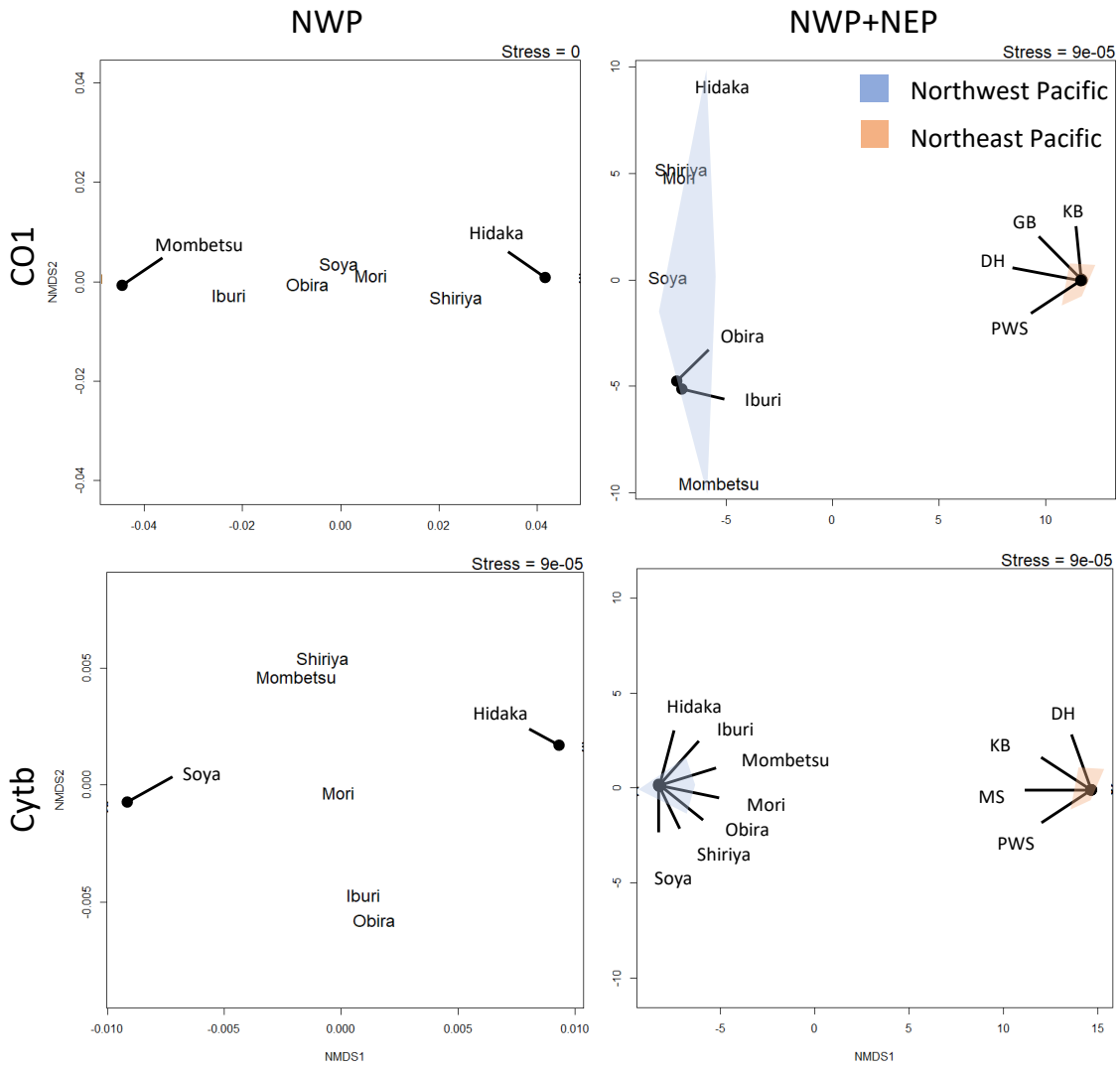


Figure 5. Nonmetric multidimensional scaling (nMDS) with Φ_{ST} scores from the Northwest Pacific (NWP) and Northeast Pacific (NEP) for each gene (COI and CYTB). Each plot was created based on the Φ_{ST} scores between sampling sites as genetic distances and the colored areas in right panels indicate the two regional groups (NWP and NEP). The stress value shown in the upper right corner of each panel represents the discrepancy between the original data and the dimensionally reduced data.

Table 4. Results of analysis of molecular variance (AMOVA) in *Enteroctopus dofleini* for the Northwest Pacific (NWP, Japan and Korea) and Northeast Pacific (NEP, Alaska and Canada) groups. F_{CT} tests the genetic variability found among NEP and NWP groups, F_{SC} tests genetic variability found among populations within NEP or NWP groups, and F_{ST} tests genetic variability found among populations.

Gene	Source of variation	d.f.	*F* statistic value	Sum of squares	Variance	% explained	p
COI	Among groups	1	F_{CT}: 0.890**	305.070	1.675	89.06	< 0.001
	Among populations within groups	18	F_{SC}: 0.060**	7.503	0.012	0.66	0.005
	Within populations	384	F_{ST}: 0.897**	74.197	0.193	10.27	< 0.001
CYTB	Among groups	1	F_{CT}: 0.865**	184.578	1.147	86.55	0.001
	Among populations within groups	12	F_{SC}: −0.014	1.447	−0.002	−0.18	0.680
	Within populations	364	F_{ST}: 0.864**	65.811	0.181	13.64	< 0.001

** $p < 0.01$

zero. Fu's F_S values were negative except for 0.107 in Mombetsu, 1.749 in Shiriya, and 0.153 in Mori, and significantly differed from zero at Soya (range: −2.508 to 1.749). The overall Fu's F_S values in NWP were negative (−7.157) and differed significantly from zero. In the NEP, both Tajima's *D* and Fu's F_S were negative (−1.562 and −1.316, respectively) and significantly differ from zero only in Tajima's *D*.

2.3.2 Cytochrome b (CYTB)

A 724 bp segment of the CYTB region was successfully sequenced in 259 of the 261 individuals (Table 1). After each sequence was trimmed to 597 bp, 9 haplotypes with 11 polymorphic sites were detected, including 3 parsimony informative sites. Haplotype EdCB_1 was the most abundant in Japan, occurring at every site and in 250 specimens, as well as in all individuals in South Korea. The seven other haplotypes found in the present study were few and closely related to EdCB_1 (Fig. 3B).

The haplotype frequencies at each sampling site are shown in Appendix B. For the CYTB region, 9 haplotypes were detected in the NWP and 10 in the NEP. Haplotype EdCB_10 appeared in Alaska as well as one individual at Hidaka, which had EdCO_9 in COI. In the newly studied samples, haplotype diversity was highest in Obira ($h \pm SD$ = 0.201 ± 0.085). In the NWP, the overall haplotype diversity ($h \pm SD$) was 0.090 ± 0.024,

ranging at sites from 0.042 to 0.201, and the overall nucleotide diversity ($\pi \pm$ SD) was 0.00002 ± 0.00010 (n = 262) (Table 1). These indices are lower than published sequence indices from the NEP ($h \pm$ SD = 0.691 ± 0.026 and $\pi \pm$ SD = 0.00150 ± 0.00010, n = 116).

In the phylogenetic tree for CYTB (Fig. 4B), The bootstrap value for the cryptic species was high which supported the independence of this clade as same as the tree for COI (Fig 4A). Within the NWP and NEP, the NEP looked like as a nested structure, but the low bootstrap value (bootstrap score 50, posterior probability 62) showed the unclear relationships of those haplotypes from the NWP and NEP.

All pairwise comparisons between the NEP and NWP showed high pairwise Φ_{ST} scores (range: 0.699 to 0.959), and 27 of the 32 comparisons showed a significant deviation from zero (Table 2). In the nMDS analysis, no groups showed correlation with specific regions or distances within the NWP population (Fig. 5). When compared with the NEP population, the NEP and NWP were separated into two distinct groups along NMDS1. Like COI, pairwise comparisons within each group showed no significant differentiations. The regional hierarchical population structure was also supported by the results of AMOVA (Table 4).

Tajima's *D* values were negative at all sites and significantly differed from zero at Mombetsu, Obira, Shiriya, and Hidaka (range: −2.002 to −1.107) (Table 1). The overall

Tajima's *D* value in the NWP was negative (−2.166) and significantly different from zero. Fu's F_S values were negative at all sites and significantly diffed from zero by −4.657 at Mombetsu, −1.602 at Soya, and −3.100 at Shiriya. The overall Fu's F_S values in the NWP were negative (−15.082) and differed significantly from zero. In the NEP, both Tajima's *D* and Fu's F_S were negative (−1.126 and −4.264, respectively) and significantly differed from zero only in Fu's F_S.

2.4 Discussion

This study showed the genetic homogeneity of *E. dofleini* among sampling sites in northern Japan. The genetic homogeneity is presumably related to dispersal during the planktonic paralarval stage, which laboratory observations suggest may last about 35–60 days (Okubo 1980, 1989). In the North Pacific, Kubodera (1991) collected *E. dofleini* paralarvae thought to be 1–2 weeks old as far as 300 miles (480 km) from shore, indicating that their dispersal can be widespread by currents. Limited population genetic differentiation is often seen in marine organisms with high dispersal potential (Palumbi 1994) and has been reported in other cephalopods with planktonic paralarvae (*e.g.*, *Architeuthis* spp., Winkelmann et al. (2013); *Octopus vulgaris* in the Mediterranean Sea, Melis et al. 2018).

Genetic variation can be used to determine stock structure, which is important for fisheries management. However, a small amount of larval mixing can easily erase the genetic signature of population structure, limiting the usefulness of genetic data for examining stock structure (Waples 1998; Begg et al. 1999). Catches of *E. dofleini* in Hokkaido are currently separated into 11 regions based on similarities in catch trends (Sano 2010) and evaluated in each region (Wakkanai Fisheries Research Institute 2022). I found no evidence of genetic structure in northern Japan to support any of these groupings, but they may have formed too little differences to discriminate using mitochondrial markers. These groupings should remain valid until markers that are more sensitive to reproductive isolation, such as microsatellites, or the use of a larger SNP library, such as RAD-seq and multiplexed ISSR genotyping by sequencing (MIG-seq), can be used to clarify the fine scale population structure. In Chapter 3, the population structure will be further examined with a greater number of SNPs using RAD-seq and microsatellite markers which have more diversity.

Genetic differentiation was found in *E. dofleini* between the NWP and NEP in haplotype network and some divergence indices, however, the phylogenetic tree did not support the branch. A plausible explanation for the genetic unclear differentiation between

the NWP and NEP is related to past fragmentation and recent population expansion on both side of North Pacific.

Genetic drift is common following population bottlenecks, and limited gene flow following the LGM could have maintained a separation of the NEP and NWP populations. During the LGM, 26,500–18,000 years ago (Kohfeld & Chase 2011; Pelto et al. 2018; Rae et al. 2020), ice sheets covered much of southern Alaska in the central part of the present distribution range for *E. dofleini* across the North Pacific, expanding across the subarctic Pacific in winter, mean annual sea surface temperatures were 5–8ºC cooler than today in the Gulf of Alaska, and the sea level was about 130 m lower (Clark & Mix 2002; Mann & Gaglioti 2024). Based on demographic analyses, it has been suggested that marine species in the North Pacific have shown various signals during the environmental changes that occurred at this time, including persistence, post-LGM recolonization, and population expansion (*e.g.*, Marko et al. 2010; Jenkins et al. 2018). Star-shape patterns in haplotype networks are typical indicators of recent population expansion (Jenkins et al. 2018), such as the haplotypes radiating from EdCO_2 in COI and from EdCB_1 in CYTB. The low nucleotide diversity and significantly negative Tajima's *D* and Fu' s *F*s within the NEP and NWP also suggest a recent population expansion. Such divergence between the NEP and NWP has presumably been maintained by the geographic separation. In Chapter 4,

demographic analysis and divergence time estimation will be conducted to further investigate the detailed historical process.

In many organisms, genetic differentiation commonly increases with geographic distance due to limited gene flow, resulting in a pattern of isolation by distance (Wright 1943). Previous studies of *E. dofleini* in Alaska suggest there are ecological differences in octopuses between the NEP and NWP (Scheel & Bisson 2012; Conrath & Conners 2014). This ecological differentiation suggests that the passive dispersal of paralarvae stage may have been too small to establish a panmictic population throughout the North Pacific Rim. I could not estimate the level of gene flow between the NEP and NWP because of a lack of samples from the central North Pacific. Considering the ecological differentiation and limitations of larval dispersal, it is unlikely that natural migration over the 7,000 km between the two regions is being maintained in the wild. However, one specimen collected at Hidaka (Japan) shared specific haplotypes in each gene (EdCO_9 in COI and EdCB_10 in CYTB) that were found otherwise only in the NEP. This suggests the possibility of past migration and survival of individuals predating the division between the two regions or human-mediated migration.

The clade comprising the cryptic species was not found in the NWP in the present study. Misawa et al. (2021) reported this cryptic species lineage from Aomori Prefecture

(northern Honshu) having the same COI sequence reported in the NEP. Toussaint et al. (2012a) suggested the distribution of the cryptic species of *E. dofleini* was limited and the population size was small. Given the likelihood that its distribution in Japan is limited and population size is small, it is possible that specimens were not captured in the regions targeted in this study.

Comparisons of genetic diversity between the NWP and NEP yielded inconclusive results. Diversities were higher in the NWP than the NEP in the COI region, but lower in the NWP than the NEP in the CYTB region. This study was unable to fully explain the reason for these conflicting results. It suggests that examination of multiple gene regions will be needed to accurately evaluate the genetic diversity.

Chapter 3: Population genetic analysis of giant Pacific octopus (*Enteroctopus dofleini*) around Japan and estimation of recent genetic exchange based on RAD-sequencing and Microsatellite analyses

3.1 Introduction

Due to advances in next-generation sequencing (NGS), other sequencing approaches are now becoming available (Wolf & Ellegren 2017). One NGS method that could be used is restriction site-associated DNA sequencing (RAD-seq). This method uses specific restriction enzymes to fragment genomic DNA samples, and size select the fragments, and sequence them to identify genetic variations (Baird et al. 2008; Davey & Blaxter 2010). It is an inexpensive technique that can genotype thousands of single nucleotide polymorphisms (SNPs) in non-model organisms that lack a reference genome (Peterson et al. 2012; Andrews et al. 2016; Theissinger et al. 2023). An increasing number of studies have used RAD-seq to assess population diversity, and SNP-based analyses have been used to study population structure in several cephalopods (*e.g.*, jumbo flying squid [*Dosidicus gigas*], Sotil & Cueva 2019; rough cranch squid [*Cranchia scabra*], jewel enope squid [*Pyroteuthis margaritifera*], and vampire squid [*Vampyroteuthis infernalis*], Timm et al. 2020; and European squid [*Loligo vulgaris*], García-Mayoral et al. 2024).

These methods have been able to reveal even subtle, weak genetic signals indicative of population structure, which is important for species with low levels of genetic differentiation because traditional methods might not reveal these differences.

Dispersal plays an important role in structuring marine populations (Chust et al. 2016), but identifying its effect on population diversity and structure is challenging especially in species with a larval phase because dispersal patterns are complex and affected by various factors. BayesAss edition 3-SNPs (BA3-SNPs) ver. 1.1.0 (Wilson & Rannala 2003; Musmann et al. 2019) is a software program that uses multi-locus genotype data and a priori parameter information to estimate recent migration rates (gene flow) and dispersal patterns between populations by performing multiple simulations. It has also been used to analyze SNP data (*e.g.*, Cheng et al. 2021; Madeira et al. 2023).

Even with the ongoing development of NGS technologies, conventional methods still have advantages in some contexts. Microsatellites are a class of tandemly repeated sequences located within genomic repetitive regions and typically exhibit high allelic diversity per site. They have long been used to estimate genetic diversity within populations and recent genetic relationships among individuals (Vieira et al. 2016). Furthermore, with the introduction of cost-effective and high-throughput multiplexing PCR

techniques and marker development methods using next-generation sequencing, microsatellites remain a highly effective tool (Guichoux et al., 2011). Considering their simplicity for genotyping, compatibility with previous data, and cost efficiency, microsatellite-based analyses continue to be a valuable approach for accumulating genetic information.

In this chapter, two methods were applied to a subset of *E. dofleini* samples from the NWP population in Japan to investigate finer-scale genetic differences and to infer gene flow among populations. The SNP data obtained through double digest RAD-seq (ddRAD-seq) using two restriction enzymes (Peterson et al. 2012, Aguirre et al. 2023) to show the genetic structures and diversity. BA3-SNPs was used to estimate migration rates (gene flow) between sample collection sites. Microsatellite analyses were also conducted to compare the population diversity with SNP results and previous research.

3.2 Materials and Methods

3.2.1 Samples

The collection of samples used in this study is described in section 2.2.1. Genomic DNA was extracted using the NucleoSpin DNA RapidLyse kit following the

manufacturer's protocol. DNA concentrations in the samples were quantified using the Gene Quant Pro RNA/DNA Calculator (ver. 4082 2.6).

A total of 32 individuals with DNA concentrations > 10 ng/μL were selected. They were collected at six sites in Hokkaido (Obira, Hidaka, Iburi, Mombetsu, Soya, and Mori) and three sites in Honshu (Asamushi, Fukui, and Shiriya) (Fig. 2, Table 1, and Appendix C). Sample numbers per site ranged from one to five (mostly four).

3.2.2 ddRAD-seq library construction

The extracted DNA was diluted in distilled water to a final volume of 20 μL and concentration of 10 ng/μL for ddRAD analyses. Due to the low concentration of DNA in one sample (Sample ID: Fukui22), a larger volume (60 μL) was used to ensure inclusion in the downstream analyses. Paired-end library construction and sequencing were performed by the Bioengineering Lab. Co., Ltd. (Kanagawa, Japan), which prepared a ddRAD library using the restriction enzymes Mspl and EcoRI, and the DNBSEQ-G400 benchtop sequencer (MGI Tech Co., Ltd.) with a 2 x 100-bp setup. A total of 28,085,491 raw reads (median 877,672 reads per sample) were obtained from the 32 samples (Appendix C).

3.2.3 Single nucleotide polymorphism (SNP) detection

Adaptor sequences and low-quality reads were trimmed using fastp ver. 0.23.4 (Chen et al. 2018). The paired-end data were trimmed in three steps:

1. The adaptor sequence, first five nucleotides, low quality reads, short reads, low mean quality reads, and long missing parts were removed.
2. In each read, overrepresented sequences were removed manually to become <1% of the read sequence.
3. All reads were adjusted to a uniform length, and low quality and short reads were removed.

In each step, the read quality and size were assessed using FastQC ver. 0.12.1 (Andrews 2010) as a quality control. In step 3, the sequence length was set to 100 bp based on the genotyping rate, and FastQC was used to check the per base sequence content, per sequence GC content, and overrepresented sequences. These reads were mapped to the reference genome assembly (accession no. GCA_027564315) using the STACKS ref_map.pl pipeline program ver. 2.65 (Catchen et al. 2013) for the paired-end ddRAD data. All 32 specimens were subjected to mapping analysis and yielded 266,434 loci, including 30,856 variant sites.

To maximize the number of SNPs and minimize the missing data, the STACKS pipeline executed each of the following filter options: population, the minimum percentage of individuals across populations required to process a locus 0.1; a maximum observed heterozygosity 0.6; and a minimum minor allele frequency 0.01. Only one SNP was randomly retained from each ddRAD locus to minimize the linkage disequilibrium (Larson et al. 2014). One specimen with much missing data (> 69%), Fukui22, was excluded. The dataset of 9,811 SNPs was filtered using PLINK ver. 1.9 (Purcell et al. 2007; Chang et al. 2015) and PCAdapt (Duforet-Frebourg et al. 2014) in R (R Core Team 2024) for downstream analysis. Two datasets, Dataset Individual (IND) and Dataset Population (POP), were prepared using different filtering conditions. For Dataset IND, strict filtering to reduce the noise of missing alleles was used to estimate the number of clusters in the population and relationships of each individual. Dataset POP included all specimens from sampling sites with more than three specimens, and the data were pooled for each sampling site. Lenient filtering allowing the inclusion of missing alleles was applied.

3.2.4 Individual-based analyses

In Dataset IND, the following filtering conditions were used in PLINK to reduce noise: minor allele frequency > 0.05 and minimum genotype rate 0.8. Outlier loci

identified through clustering with q-values < 0.01 using the PCAdapt package in R (Privé et al. 2020) were removed from the dataset as non-neutral SNPs. The most likely value of *K* (number of Principal Components) was determined based on the cumulative explained variance > 80%. The proportion of missing loci in each individual was confirmed using PLINK. The sample Soya203 was excluded because it had a high (> 20%) missing rate. After filtering, 1,526 unlinked nuclear SNPs from 30 individuals from nine sites were retained in Dataset IND. The total genotyping rate was 93.1%. Details of the filtering effects are presented in Appendix C.

To explore genetic variation and potential groupings, four individual-based analyses were conducted with Dataset IND. Initially, I obtained a covariance matrix and performed individual-level principal component analysis (PCA) using PCAdapt. Neighbor-Net networks analyses were constructed based on p-distances using SplitsTree4 ver. 4.19.2 (Huson & Bryant 2006) to show the relationships between sequences as a network. To estimate shared ancestry, individual admixture proportions were estimated using the program STRUCTURE (Pritchard et al. 2000). Ten independent runs were performed for *K* values (*i.e.*, number of supposed ancestor) ranging from 1 to 10, with each run using 10 iterations for the initial analyses, and the MCMC was increased to 500,000. The most likely value of *K* was estimated using the mean LnP (*K*) and Δ*K* method described by

Evanno et al. (2005) using StructureHarvester ver. 0.6.94 (Earl & von Holdt 2011), and the CLUMPAK Server was used to combine and visualize the STRUCTURE results (Kopelman et al. 2015).

Discriminant analysis of principal components (DAPC) was performed to identify and describe clusters of genetically related individuals without a priori information using the adegenet package in R (Jombart & Ahmed 2011). In these datasets, the proper number of clusters could not be selected from the BIC, so $K = 2$ to 10 clusters were tested to compare with the STRUCTURE analyses. The proper number of principal components was determined based on the cumulative explained variance > 80%. The validity of the proportions of successful assignments was shown in an assignment plot, and the clustering was shown in a scatter plot and component plot.

3.2.5 Population-based analyses

Willing et al. (2012) showed that for SNP numbers > 1,000, RAD-seq can provide accurate estimates of population genetic differentiation with small sample sizes of 2–6 per site. In Dataset POP, I excluded the two populations with few specimens (Asamushi and Fukui). Soya203 had a high (> 30%) missing rate and was also excluded. Next, the following filtering conditions were applied using PLINK for the downstream population

analyses: minor allele frequency > 0.05 and minimum genotype rate 0.6. Outlier loci with q-values < 0.01 using PCAdapt were considered outlier SNPs and excluded. The most likely value of *K* was determined in the same way as for Dataset IND. The proportion of missing loci in each individual was confirmed using PLINK. In Dataset POP, 2,127 SNPs from 28 individuals were retained, and the total genotyping rate was 83.7%. Details of the filtering effects are presented in Appendix C.

Seven diversity indices were calculated at each collection site using GenAIEx ver. 6.5 (Peakall & Smouse 2012): allele number (N_A), total number of alleles; effective allele number (N_E), total number of alleles conformed Hardy-Weinberg equilibrium; Shannon index (I), the diversity of alleles based on the frequency of each alleles; observed heterozygosity (H_O), the observed number of heterozygous loci; expected heterozygosity (H_E), calculated from the frequency of alleles; unbiased expected heterozygosity (uH_E), heterozygosity calibrated by sample size; and fixation index (F). Principal Coordinate Analysis (PCoA) to compare the combination of SNPs and pairwise F_{ST} to show the genetic distance were also performed for each collection site using GenAIEx. To test the correlation between genetic and geographical distances, the Mantel test was used with 999 permutations in R using the vegan package. The shortest geographical distance (km) between locations along the coastline was measured with ArcGIS Desktop ver. 10.

Pairwise genetic distances were calculated as F_{ST} / (1 − F_{ST}) as recommended by Rousset (1997).

3.2.6 Migration analyses

BA3-SNPs was used to estimate the pairwise directional gene flow (migration rate) between sampling sites using Dataset POP. The analyses were performed using BA3-SNPs ver. 1.1 (Wilson & Rannala 2003; Mussmann et al. 2019). BA3-SNPs uses a Bayesian MCMC approach to estimate asymmetric recent migration rates, which are the proportion of each population with migrant ancestry. A total of 10^8 iterations were run with a burn-in of 10^6 and sampling interval of 1,000. The mixing parameters for the allelic frequency, inbreeding coefficient, and migration rate were adjusted for an acceptance rate of 20–60% (Wilson & Rannala 2003). In this study, these parameters were set as follows: migration rate = 0.5, allele frequencies = 0.7, and inbreeding coefficients = 0.2. The acceptance rates for the hypothesis ranged between 30 and 60%, making these values appropriate for the analysis. The convergence of all parameters was accessed using Tracer ver. 1.6 to confirm the ESS > 200.

3.2.7 Microsatellite analyses

A total of 18 microsatellite markers were amplified with primers developed by Toussaint et al. (2012b). These markers were tested using the KOD One PCR Master Mix (Toyobo, Osaka, Japan) and PCR amplified product was confirmed using 3% agarose gel electrophoresis. The PCR reaction mix was prepared following the manufacturer's protocol. Of these, seven markers that could amplify samples were selected for following PCR protocol (marker details are presented in Appendix D). PCR amplification was performed using the KOD One PCR Master Mix and KOD-Multi & Epi-^{TM} DNA polymerase (Toyobo, Osaka, Japan). The PCR reaction mix was prepared following the manufacturer's protocol with each tailed universal fluorescent primer (6-FAM, VIC, NED or PET). The following thermal conditions were used: initial denaturation at 94℃ for 2 min followed by 30 cycles of denaturation at 98℃ for 10 s, annealing at 55 ℃ for 20 s and extension at 68℃ for 15 s, followed by 5 cycles of denaturation at 98℃ for 5 s, annealing at 45 or 48 ℃ for 15 s and extension at 68℃ for 10 s, followed by a final extension at 68℃ for 2 min. The PCR products were visualized using agarose gel electrophoresis and analyzed using ABI PRISM 3130xl Genetic Analyzer (Applied Biosystems).

Fragment analysis was conducted using Thermo Fisher Connect Platform (Thermo Fisher Scientific, Waltham, MA, USA). Alleles for a locus were recoded as the length of

fragments. Due to the difficulty in comparing SNP and microsatellite results (Fischer et al. 2017), I conducted only part of the analyses of SNP data to compare the diversity pattern and trends. The seven diversity indices listed above (N_A, N_E, I, H_O, H_E, uH_E, and F) were also calculated at each collection site, and PCoA was performed using GenAIEx ver. 6.5.

3.3 Results

3.3.1 Population structure

The first four analyses (PCA, STRUCTURE, DAPC, and Neighbor-Net) were based on genotype likelihoods and verified the potential population structure. In PCA, K = 1 was selected as the most appropriate for the false discovery rate based of the cumulative explained variance (Appendix E). The PCA using Dataset IND found no population structure in Hokkaido with PC1 and PC2 (Fig. 6A), or with PC3 and PC4 (Fig. 6C). One outlier, Fukui 6, was observed based on PC2, but as only one specimen was available from Fukui, it was not possible to determine if it formed a distinct cluster. In the PCA scatterplot colored with the mitochondrial haplotypes in COI (Fig. 6B), the ellipsoids about the two main haplotypes (1 and 2) showed high overlap. The haplotype also showed extensive overlap in PC3 and PC4, making it difficult to distinguish the two haplotypes (Fig. 6D). Similar results (*i.e.*, no clustering observed by haplotype or sampling site) were found in

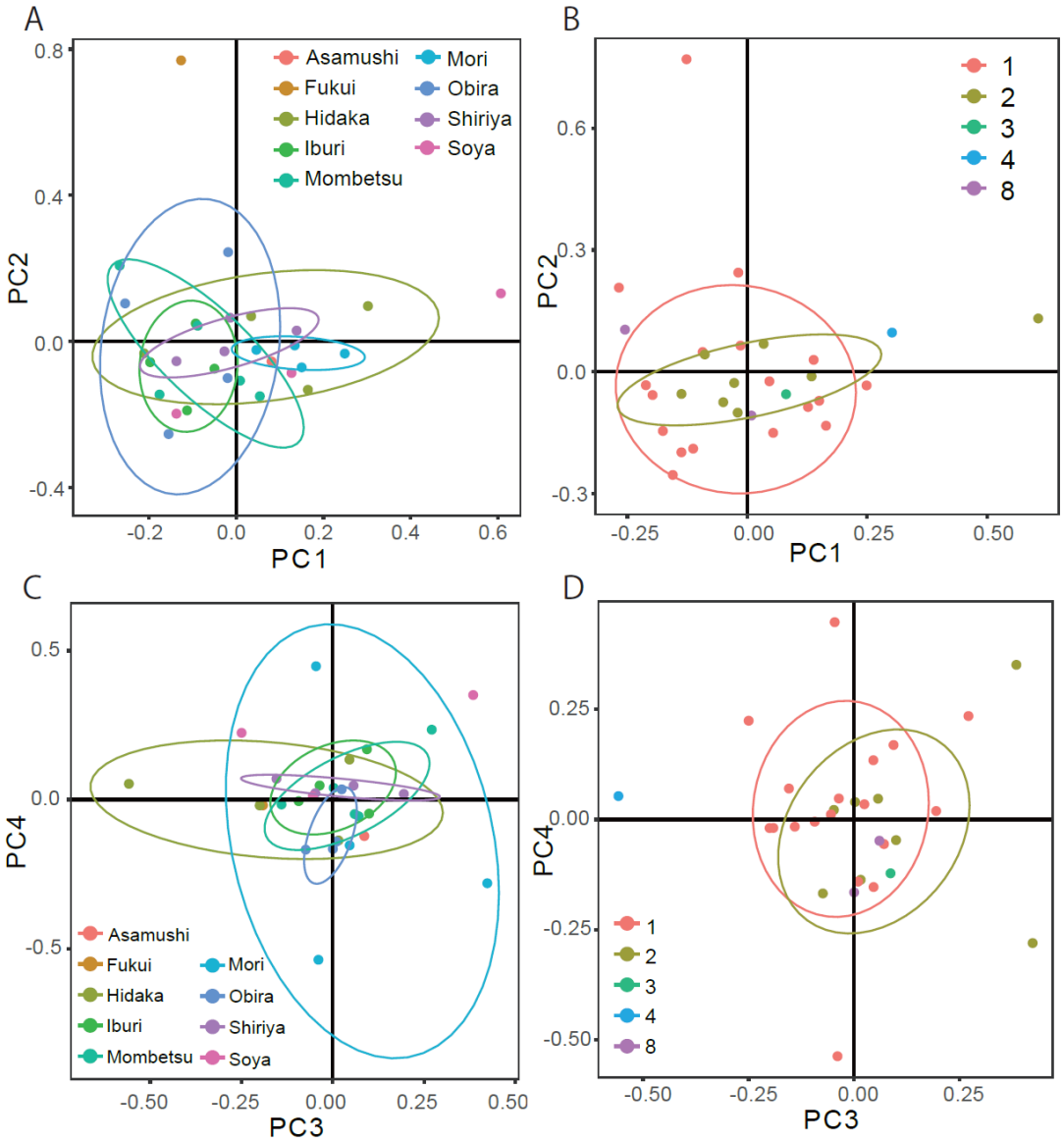


Figure 6. Principal component analysis (PCA) of population structure using Dataset IND. A and B; Scatter plots of PC1 vs PC2 based on Dataset IND. C and D; Scatter plots of PC3 vs PC4 based on Dataset IND. Colors represent sampling sites in A and C and haplotypes from the Chapter 2 in B and D.

the Neighbor-Net algorithm. Neighbor-net analysis revealed several unclear clusters, but they did not correlate with geographical distances (Fig. 7).

In the STRUCTURE analysis, based on Mean LnP (*K*) and Delta *K*, *K* = 3 was selected as the most likely *K* value (Figs. 8A, B). Using this value, some sites formed a cluster but the boundary was unclear. For all values of *K*, the population data did not show discriminable patterns of separation into distinct groups (Fig. 8C). To compare the results with those using STRUCTURE, DAPC was performed with *K* values ranging from 2 to 10 (Appendices F, G). However, none of the clusters were related to geographical distances.

3.3.2 Genetic diversity and population analyses

The genetic diversity indices from the population-based analyses using Dataset POP are summarized in Table 5. The Shannon diversity index (*I*) ranged from 1.273 in Soya to 1.354 in Mombetsu. Values were similar across locations. All indices showed similar trends; values tended to be highest in Hidaka and Mombetsu and lowest in Soya. With respect to Soya, the small number of individuals included in the sample may have influenced these results. The fixation index (*F*) was positive and near 0 at all sites. The value was lowest 0.219 in Soya and highest 0.259 in Hidaka.

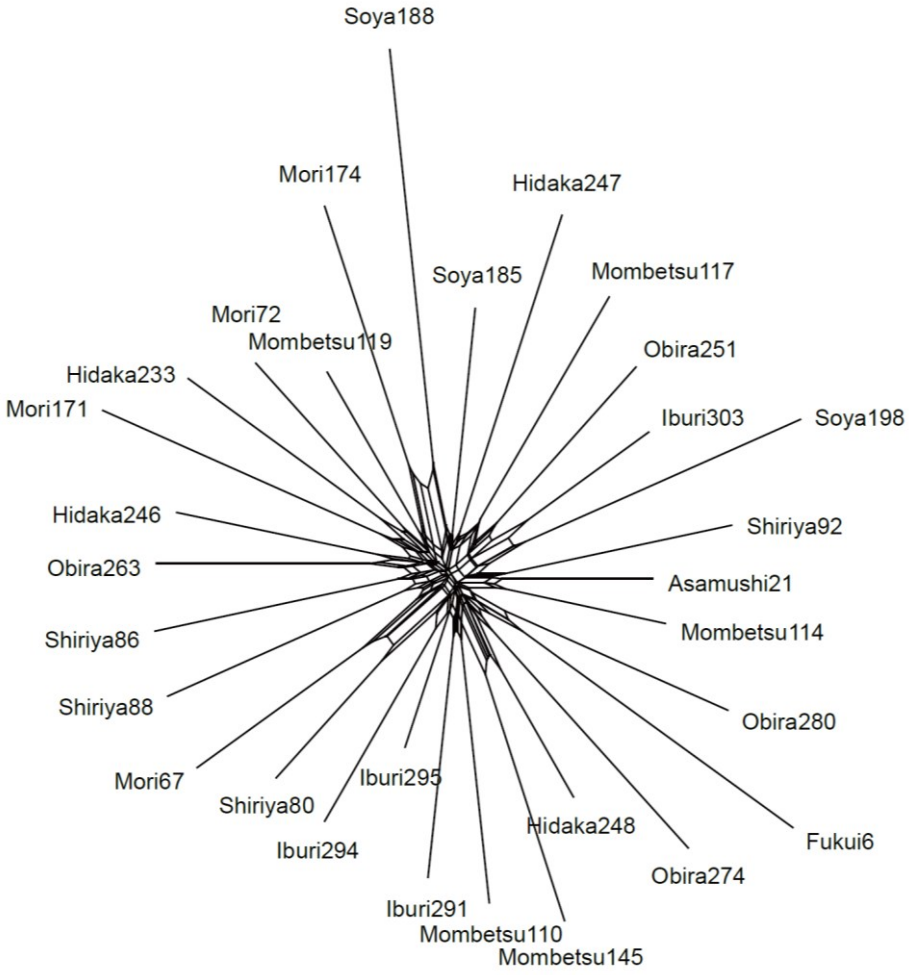


Figure 7. Neighbor-Net network using Dataset IND. Node tips indicate the sampling site and sample number.

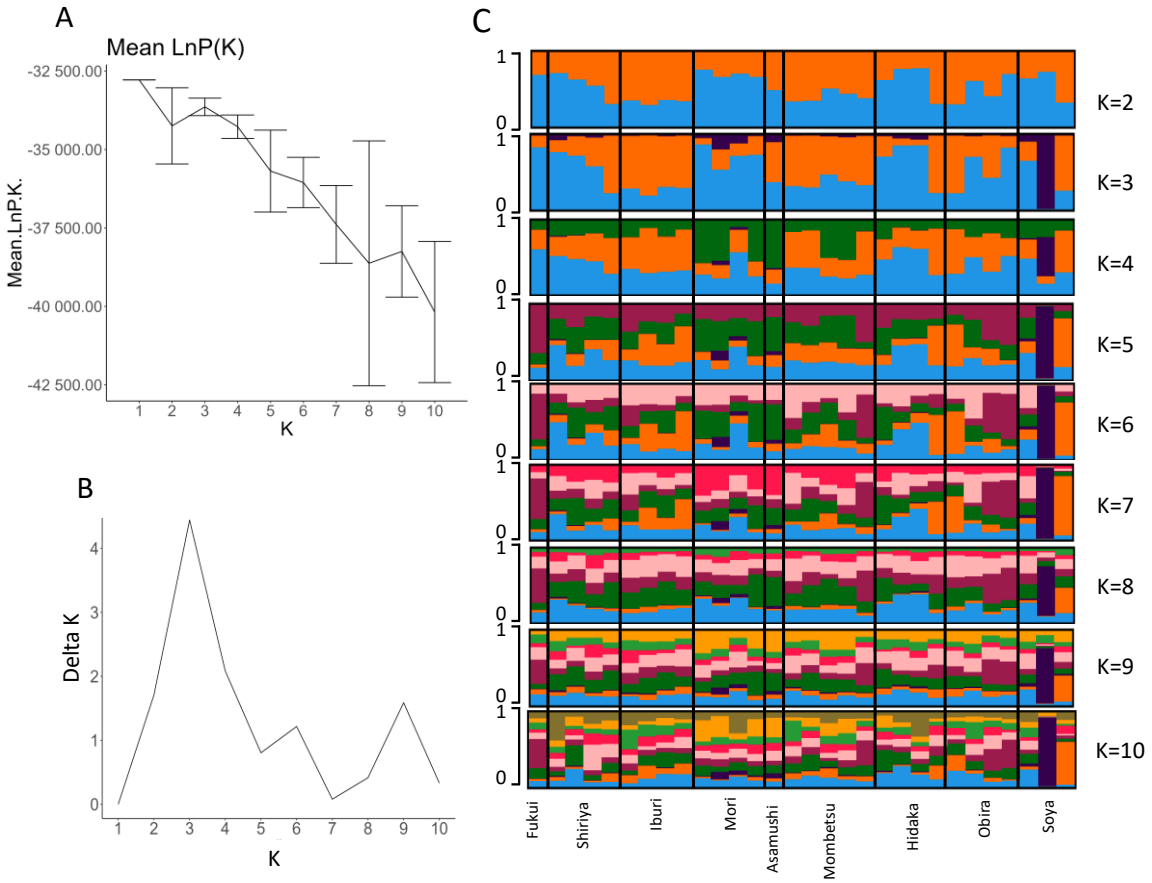


Figure 8. Results of STRUCTURE analysis. A and B: Analysis to determine the most likely *K* value. C: Bar plot of STRUCTURE analysis using Dataset IND for *K* values from 2 to 10. Each vertical bar indicates an individual, and the y-axis represents the admixture proportion for potential ancestral populations.

Table 5. Sampling sites and genetic diversity of *Enteroctopus dofleini* based on ddRAD-seq. Population genetic analysis was conducted excluding Fukui and Asamushi, where sample numbers were small.

Sites	n	N_A	N_E	I	H_O	H_E	uH_E	F	Polymorphic loci (%)
Hidaka	4	3.378 ± 0.018	1.619 ± 0.011	1.352 ± 0.008	0.326 ± 0.006	0.246 ± 0.006	0.215 ± 0.004	0.259 ± 0.005	62.2%
Iburi	4	3.244 ± 0.020	1.586 ± 0.011	1.331 ± 0.008	0.311 ± 0.006	0.236 ± 0.006	0.205 ± 0.004	0.249 ± 0.005	59.2%
Mombetsu	5	4.245 ± 0.021	1.674 ± 0.010	1.354 ± 0.007	0.336 ± 0.006	0.236 ± 0.005	0.219 ± 0.004	0.251 ± 0.005	67.4%
Mori	4	3.493 ± 0.016	1.617 ± 0.011	1.345 ± 0.007	0.321 ± 0.006	0.236 ± 0.006	0.211 ± 0.004	0.250 ± 0.005	61.8%
Obira	4	3.329 ± 0.019	1.599 ± 0.011	1.340 ± 0.008	0.317 ± 0.006	0.230 ± 0.006	0.209 ± 0.004	0.250 ± 0.005	60.4%
Shiriya	4	3.283 ± 0.020	1.595 ± 0.011	1.342 ± 0.008	0.317 ± 0.006	0.234 ± 0.006	0.210 ± 0.004	0.252 ± 0.005	59.9%
Soya	3	2.464 ± 0.016	1.435 ± 0.011	1.273 ± 0.008	0.252 ± 0.006	0.207 ± 0.006	0.169 ± 0.004	0.219 ± 0.006	44.9%
TOTAL	28	3.348 ± 0.008	1.589 ± 0.004	1.334 ± 0.003	0.311 ± 0.002	0.232 ± 0.002	0.205 ± 0.002	0.247 ± 0.002	59.4%

n, number of examined individuals: N_A, number of alleles: N_E, number of effective alleles: I, Shannon index, H_O, observed heterozygosity: H_E, expected heterozygosity: uH_E, unbiased expected heterozygosity: F, fixation index,

Table 6. Sampling sites and genetic diversity of *Enteroctopus dofleini* based on the seven microsatellite markers. Population genetic analysis was conducted excluding Fukui and Asamushi, where sample numbers were small.

Sites	n	N_A	N_E	I	H_O	H_E	uH_E	F	Polymorphic loci (%)
Hidaka	4	3.000 ± 0.378	2.028 ± 0.354	0.788 ± 0.144	0.429 ± 0.090	0.438 ± 0.071	0.500 ± 0.081	-0.001 ± 0.137	100.00%
Iburi	4	2.000 ± 0.378	1.649 ± 0.310	0.462 ± 0.185	0.286 ± 0.115	0.272 ± 0.111	0.311 ± 0.127	-0.084 ± 0.082	57.14%
Mombetsu	5	2.571 ± 0.481	1.964 ± 0.355	0.669 ± 0.184	0.264 ± 0.071	0.388 ± 0.100	0.433 ± 0.111	0.240 ± 0.123	85.71%
Mori	4	2.429 ± 0.571	2.043 ± 0.467	0.605 ± 0.246	0.214 ± 0.101	0.330 ± 0.133	0.378 ± 0.152	0.271 ± 0.160	57.14%
Obira	4	3.571 ± 0.369	2.505 ± 0.322	1.016 ± 0.139	0.524 ± 0.097	0.549 ± 0.069	0.632 ± 0.079	0.038 ± 0.121	100.00%
Shiriya	4	3.000 ± 0.756	2.648 ± 0.664	0.820 ± 0.271	0.393 ± 0.143	0.438 ± 0.132	0.500 ± 0.150	0.149 ± 0.190	71.43%
Soya	3	2.286 ± 0.360	1.853 ± 0.321	0.622 ± 0.155	0.357 ± 0.077	0.379 ± 0.086	0.462 ± 0.103	-0.032 ± 0.141	85.71%
TOTAL	28	2.694 ± 0.187	2.098 ± 0.155	0.712 ± 0.073	0.352 ± 0.039	0.399 ± 0.038	0.459 ± 0.044	0.077 ± 0.051	100.00%

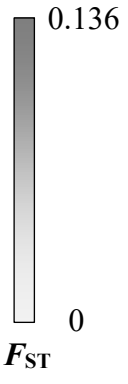
n, number of examined individuals: N_A, number of alleles: N_E, number of effective alleles: I, Shannon index, H_O, observed heterozygosity: H_E, expected heterozygosity: uH_E, unbiased expected heterozygosity: F, fixation index,

In the microsatellite analyses (Table 6), values were similar across locations and all indices showed similar trends except for Soya. N_E values were similar to those value obtained from the ddRAD-seq results, but N_A and I values were slightly lower. In I, He, and uHe, higher genetic diversity was observed in Obira, Shiriya, and Hidaka, whereas lower diversity was found in Iburi and Mori within Funka Bay. The inbreeding coefficient (F) showed negative values in Hidaka, Iburi, and Soya. These marker sets were slightly different from Toussaint et al. (2012b), so it was difficult to compare with the results of previous study.

The pairwise F_{ST} values ranged from 0.083 to 0.136, with relatively high values in Soya (Table 7). No significant correlation between $F_{ST} / (1-F_{ST})$ and geographic distance was detected by the Mantel test (Mantel r = −0.010, p = 0.452) (Fig. 9). The six plots clustered in the upper part of the graph because of the relatively high F_{ST} values between Soya and the other sites. In the PCoA, PC1 and PC2 explained 72.5% of the variance (Fig. 10). This PCoA plot represents genetic distances among sampling sites. Obira, Soya, and Mombetsu (northern Hokkaido) were closely clustered along both PC1 and PC2, whereas no clear geographic clustering was observed among the other sites. In the microsatellite analyses (Fig. 11), PC1 and PC2 explained 100% of the variance in the PCoA, and no clear geographic clustering was observed among the other sites.

Table 7. Pairwise F_{ST} scores (below diagonal) for all population pairs excluding Asamushi and Fukui. The black-to-white gradient corresponds to the change from higher to lower values

	Hidaka	Iburi	Mombetsu	Mori	Obira	Shiriya
Hidaka						
Iburi	0.104					
Mombetsu	0.084	0.094				
Mori	0.090	0.103	0.083			
Obira	0.095	0.109	0.085	0.098		
Shiriya	0.097	0.110	0.093	0.098	0.107	
Soya	0.126	0.136	0.113	0.125	0.133	0.134



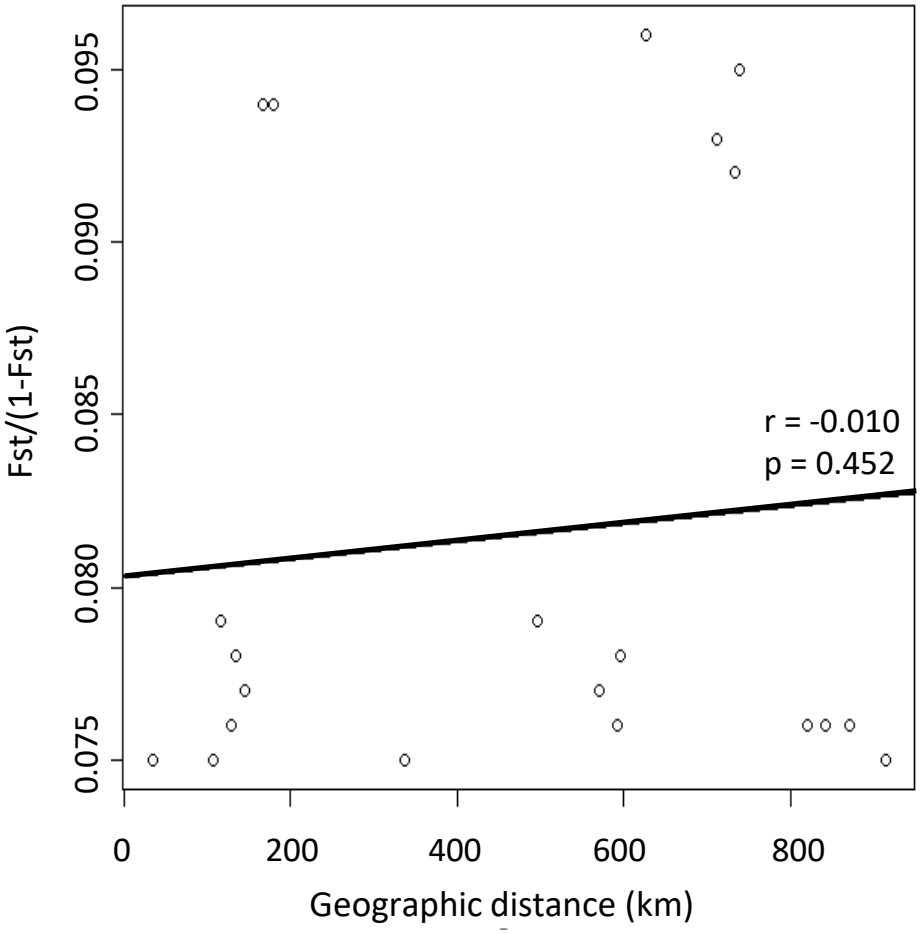


Figure 9. Relationship between genetic distance (F_{ST} / (1 – F_{ST})) and geographic distance from Mantel tests. Circles represent individual samples.

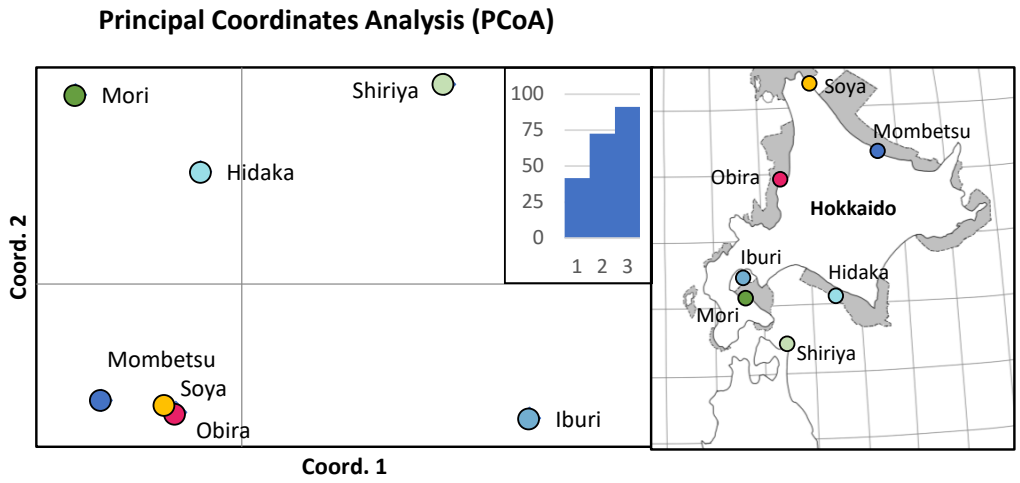


Figure 10. Principal coordinates analysis (PCoA) of Dataset POP. Bar plot (top center) shows the cumulative explained variance for each coordinate, and the center panel shows the scatter plot of PCoA. Right panel shows sampling sites for this dataset.

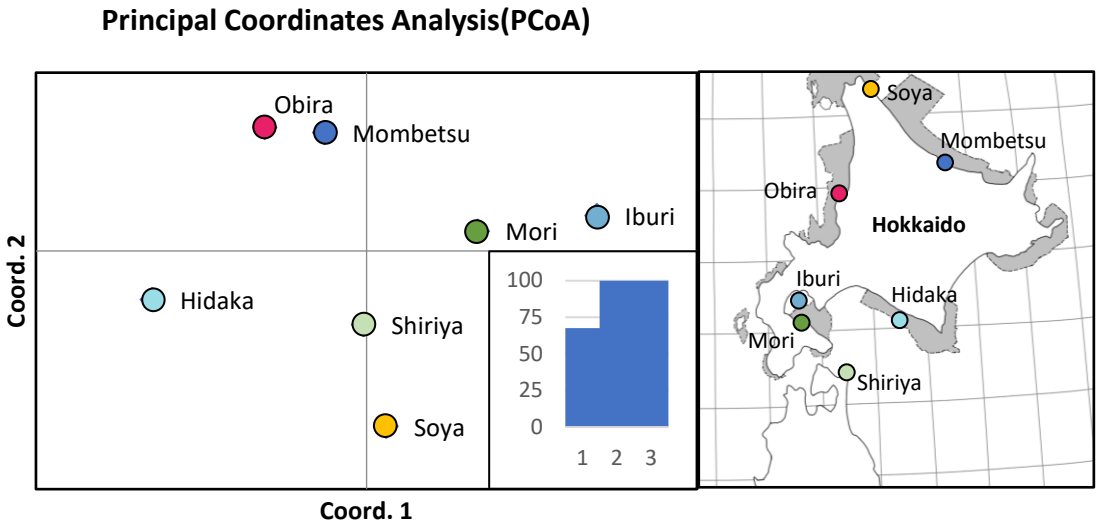


Figure 11. Principal coordinates analysis (PCoA) of microsatellite. Bar plot (bottom center) shows the cumulative explained variance for each coordinate, and the center panel shows the scatter plot of PCoA. Right panel shows sampling sites for this dataset.

3.3.3 Estimated migration

As a result of the analysis, the mean migration rate was estimated to be 0.032 ± 0.004 (mean ± SD). The mean self-recruitment rate was 0.810 ± 0.014 (Table 8). The inbreeding coefficients (F_{IS}) were similar at all sites and ranged from 0.025 to 0.123.

3.4 Discussion

The results of the SNP and microsatellite analyses of *E. dofleini* in Japan showed a close relationship and random mating among the different sampling sites apart from 1,300 km (Fukui and Mombetsu) and 900 km (Hidaka and Mombetsu). These results suggest the idea that the population functions like a single panmictic unit. In marine organisms, genetic diversity and population genetic structure are affected by larval dispersal and ocean currents (*e.g.*, White et al. 2010; Griffiths et al. 2021). Cephalopods display two major developmental modes: planktonic and benthic (Villanueva et al. 2016). The mode of hatching can have a significant impact on population structure in octopuses (*e.g.*, *Hapalochlaena maculosa*: Morse et al. 2018; *Octopus insularis*: Bein et al. 2023). Because the short-term movement range of this species has been reported to be limited (*e.g.*, 50 m² over two months [Rigby & Sakurai 2001]; 60 m over two months [Scheel & Bisson 2012]; 5 km over 200 days [Joh et al. 2017]), the observed genetic connectivity

Table 8. Estimated migration rates of *Enteroctopus dofleini* between sampling sites using BA3-SNPs for Dataset POP.

	Source							Inbreeding coefficient (F_{IS})
Sink	Hidaka	Iburi	Mombetsu	Mori	Obira	Shiriya	Soya	
Hidaka	0.812	0.030	0.036	0.030	0.030	0.031	0.031	0.045
Iburi	0.030	0.813	0.035	0.030	0.030	0.031	0.031	0.037
Mombetsu	0.028	0.028	0.834	0.028	0.028	0.028	0.028	0.025
Mori	0.030	0.031	0.035	0.814	0.030	0.030	0.030	0.066
Obira	0.031	0.031	0.053	0.030	0.795	0.030	0.030	0.077
Shiriya	0.031	0.030	0.036	0.030	0.030	0.813	0.031	0.061
Soya	0.033	0.033	0.043	0.033	0.034	0.033	0.790	0.123

over distances exceeding 800 km is likely maintained by larval dispersal.

Studies on marine mollusks around Japan have revealed diverse population structures (*e.g.*, Yamada et al. 2014; Kawamura et al. 2017; Yamazaki et al. 2017; Takeuchi et al. 2020). But many populations in the Sea of Japan are genetically relatively homogeneous. The current system is strongly influenced by the Tsushima Warm Current, which flows northeastward along the west coast of Japan, and presumably maintains north-south connectivity (Kojima et al. 2014; Azuma & Chiba 2017). This could explain the genetic similarity seen in samples from Fukui and Hokkaido located 800 to 1,300 km apart, as well as between locations in the Pacific Ocean and Sea of Okhotsk, and suggest that the Tsugaru Current and Soya Current help maintain connectivity around Hokkaido. However, for certain regions (*e.g.*, Fukui), the long distance and small sample size may have resulted in insufficient detection power, leaving room for further discussion.

No populations corresponding to the 11 geographic divisions proposed by Sano (2010) were identified both the microsatellite and ddRAD-seq analyses. These divisions were based on the catch volume and annual mean catch, and it was hypothesized that the areas between the divisions likely consisted of sandy or sloped terrains, which are thought to be unsuitable for the nesting of settled *E. dofleini*. Therefore, while the regions suggested by Sano (2010) may reflect the movement or habitat range of individuals after

settlement, they may not be suitable for inferring population structure influenced by gene flow during the larval stage.

Connectivity could also be reinforced through the long-term movement of settled individuals. There is little evidence that *E. dofleini* undergoes large scale migrations like some squids (40 km over 200 days, Noro & Sakurai 2012; Joh et al. 2017), but *E. dofleini* has relatively a long lifespan (4–5 years) and large body size (>20 kg, Bower et al. 2024), which could enhance its post-settlement mobility. In addition, multiple paternity has been reported in *E. dofleini* from the NEP based on egg samples (Larson et al. 2015). Since females of this species are capable of storing multiple spermatophores, long-distance movement by either sex could potentially enhance gene flow. This could result in greater connectivity between sites compared to bivalves and other shorter-lived cephalopods (Yamazaki et al. 2017; Melis et al. 2018).

The weak genetic structure and low inbreeding coefficients (F_{IS}: 0.025 to 0.123) in the BA3-SNPs analysis indicate that the observed heterozygosity was greater than the expected heterozygosity within each site, which suggests widespread gene flow among populations. However, this result appears to contradict the high self-recruitment rates shown by BA3-SNPs based on genetic differences among sites. According to Slatkin (1985), even rare long-distance dispersal events can counteract genetic drift. This implies

that a small number of individuals dispersing through larval movement may be sufficient to maintain genetic homogeneity across sites.

Alternatively, the apparent contradiction may stem from the low genetic diversity of SNPs included in the dataset. It has been suggested that when SNP numbers are limited or genetic differentiation among populations is minimal, assignment methods tend to overestimate the self-recruitment (Wilson & Rannala, 2003, Mussmann et al. 2019). The presence of rare minor alleles and a high proportion of missing data can also influence the results. However, similar trends were observed even in other dataset under more strict filtering conditions (*e. g.*, n = 28, minor allele frequency threshold of 0.1 or 0.07; minimum genotype rate 0.9) (data not shown). Consequently, BA3-SNPs may have chosen a very low migration rate because of the low genetic diversity despite the initial assumption of a higher migration rate, leading to an inflated estimate of the self-recruitment.

Another possible factor in the differing results between analyses are variations in the duration and frequency of data collection. F-statistics reflect relatively long-term genetic structure, while BayesAss estimates recent migration events, typically within approximately five generations. Palsbøll et al. (2007) pointed out that in populations that have low genetic diversity, using indicators of recent gene flow may be more suitable for conservation planning than relying on metrics that reflect historical connectivity. Given

that *E. dofleini* may have undergone population expansion relatively recently (discussed in Chapter 4), the observed differences between these indicators could be attributed to their differences in the time scales each method reflects.

The results of this study show that genetic analyses using both mitochondrial data (Chapter 2) and microsatellite analyses and high-resolution SNP analyses (this chapter) reveal no clear population structure in *E. dofleini* around Japan, suggesting significant gene flow across the region. This means that changes population size may not be driven by isolated regional dynamics, but may be more synchronized across broad areas. This information can be used to inform future conservation efforts and promote the sustainable management of this resource.

Chapter 4: Population dynamics and phylogenetic analysis of giant Pacific octopus (*Enteroctopus dofleini*) in the Northeast Pacific and Northwest Pacific

4.1 Introduction

The phylogeny, population structure, and population size of marine species are influenced by ocean currents and geographical barriers such as isthmuses, which both affect gene flow between populations, as well as by global climate change. The results of Chapter 2 showed that there are small genetic differences in the giant Pacific octopus (*Enteroctopus dofleini*) between the NEP and NWP, and that there may have been a recent population expansion in both regions. The genetic differentiation and temporal dynamics of the population expansion remain unclear, and it is uncertain if either might have been affected by climatic events. In the North Pacific, some species are thought to have experienced population fragmentations or range contraction during the LGM and subsequent population expansions (*e.g.*, Pacific herring [*Clupea pallasii*]: Liu et al. 2012; Chilean dolphin [*Cephalorhynchus eutropia*]: Pérez-Alvare et al. 2016). In *E. dofleini'*s closely related congener *E. megalocyathus*, which is distributed at the southern tip of South America, low genetic diversity and a closely related population expansion suggest a post-glacial recolonization may have occurred (Pliego-Cárdenas et al. 2022).

The giant Pacific octopus (*E. dofleini*) belongs to the family *Enteroctopodidae*, which comprises four genera: *Enteroctopus*, *Muusoctopus*, *Sasakiopus,* and *Vulcanoctopus* (Strugnell et al. 2014). *E. dofleini* was first described as *Polypus dofleini* by Wülker (1910) using a specimen collected in southern Hokkaido, but its classification has since undergone several revisions (Pickford 1964; Hochberg 1998; Norman et al. 2016). The current classification within *Enteroctopus* is consistent with genetic data from other large North Pacific octopuses that were also reclassified as *Enteroctopus* (Hochberg 1998; Bower et al. 2024). The genus *Enteroctopus* comprises four large species that occur in coastal to deep waters. In addition to *E. dofleini*, they include *E. megalocyathus* in the southeast Pacific and southwest Atlantic (originally described in Gould 1852), *E. magnificus* in southern America (Villanueva et al. 1991), and *E. zealandicu*s in New Zealand (originally described in Benham 1944, basis of record in Norman & Hochberg 2005). In the closely related genus *Muusoctopus*, a time-calibrated phylogenetic reconstruction study based on the fossil record with some genes suggested that speciation within the genus was related with the formation of the Isthmus of Panama (Lima et al. 2020b). In the genus *Enteroctopus*, phylogenetic studies are limited. Although recent diversification and dispersal have been suggested for *E. megalocyathus* and *E. zealandicus* (Pliego-Cárdenas et al. 2022), the

evolutionary history of the genus and the possible effects of climatic events remain unclear.

In a study of *Enteroctopus* in the NEP, Toussaint et al. (2012a) provided evidence of the existence of an undescribed and previously unrecognized cryptic *Enteroctopus* species in Prince William Sound, Alaska, that is allied to, but distinct from *E. dofleini*. Cryptic speciation can occur in organisms that evolve in extreme environments, which can sometimes limit the extent of morphological changes they undergo (called "morphological stasis"), where their appearance remains relatively stable even as genetic differences accumulate (Bickford et al. 2007). Well-resolved phylogenetic trees are essential tools, not only for exploring environmental events that influence diversification (Lima et al. 2020b) but also for identifying key traits in evolutionary processes (Oellermann et al. 2015).

In this chapter, I estimate the historical demographic change of *E. dofleini* in the North Pacific and the time of the genetic separation between populations in the Northeast Pacific and Northwest Pacific using demographic analyses. A phylogenetic time-calibrated tree was also reconstructed using fossil records of congeners within the genus *Enteroctopus* to examine the relation between major climatic event and the divergence of the genus.

4.2 Material and Methods

4.2.1 Demographic analysis

Concatenate sequences of COI and CYTB from NWP analyzed in Chapter 2 were used to conduct mismatch distribution analysis to infer past demographic events. In the NEP, the concatenate sequences were could not be constructed because individual haplotypes were not shown in Toussaint et al. (2012a). Therefore, mismatch distribution analyses were conducted separately for COI and CYTB. This made comparison between the NWP and NEP difficult, so demographic analysis was performed on the NEP data to provide a reference in the comparison of the two regions. These analyses were performed using Arlequin ver. 3.5.1.2 (Excoffier & Lischer 2010) with a generalized least-square approach to estimate parameters associated with a sudden population expansion. The validity of using a sudden expansion model was determined based on the sum of squared deviations (SSD) and the Harpending's raggedness index (Harpending et al. 1993). Statistical significance was estimated with 2,000 permutations using Arlequin. The time since the occurrence of the population expansion was estimated from $t = \tau/(2\mu k)$, where τ = the expansion time estimator, μ = the mutation rate, and k = the sequence length (COI: 517 bp, CYTB: 597 bp). In both geographic regions, I used a nucleotide divergence rate of 0.02 substitutions per site per million years (Myr) based on a previous study using the

mitochondrial COI region in *Octopus mimus* (Pardo-Gandarillas et al. 2017). To evaluate the uncertainty of the expansion estimates, 95% confidence intervals (CI) were calculated using a parametric bootstrap procedure (2,000 replicates).

Bayesian skyline plots (BSP) were used to reconstruct population fluctuations based on concatenated sequence data from observed and simulated sequences using BEAST ver. 2.7.4 (Drummond & Rambaut 2007). The HKY model was selected based on the BIC calculated using MEGA software (ver. 7.0) (Kumar et al. 2016) and Kakusan4 software (Tanabe 2007), empirical base frequencies, and a strict clock rate ($\mu = 0.02$ Myr) prior to running 50 million steps sampled every 1,000 steps, with the first 5 million steps discarded as burn-in. These runs were repeated eight times and the log files were combined using LogCombiner. The convergence of combined log data was verified by both visually inspecting the MCMC traces in Tracer v1.6 to confirm the ESS > 200. The exact generation time (g) for *E. dofleini* is unclear (3–5 years; Noro & Sakurai 2012), so the y-axis was presented as effective population size (Ne) and generations (T), *i.e.*, NeT.

4.2.2 Estimation of divergence times

In the initial analysis to construct a reference tree, the complete mitochondrial sequences from 11 species were selected based on calibration points (Appendix H). These

sequences were downloaded from the National Center for Biotechnology Information (NCBI), and protein-coding regions were partitioned using MEGA. The complete mitogenomes were fragmented as 13 protein coding regions (COI, COII, COIII, CYTB, ATP6, ATP8, ND1–ND6, and ND4L), and unalignable regions were removed using Gblocks software ver. 0.91b in the default parameters (Castresana 2000), and merged in each region. Several sequences of the 13 protein coding regions for three closely related species (*E. megalocyathus*, *E. zealandicus*, and *Vulcanoctopus hydrothermalis*) were added in the partitioned sequence files from the NCBI (details are shown in Appendix H). For the NWP region, NEP region, and cryptic species, a representative sequence was selected for each mitochondrial region (COI: 1535 bp; CYTB: 1138 bp) based on the most dominant haplotype in the haplotype network and added to the partitioned alignment files. For COI, the haplotypes were EdCO_1 (NWP), EdCO_9 (NEP), and EdCO_16 (cryptic species). For CYTB, the haplotypes were EdCB_1 (NWP), EdCB_15 (NEP), and EdCB_24 (cryptic species) (Appendix B). A phylogenetic tree was constructed using IQ-TREE 2 software (Minh et al. 2020), with the partitioned dataset and substitution model selection. The following best fit models were also detected separately using IQ-TREE 2 for each protein-coding region: GTR+F+G4 for ATP6; GTR+F+I+G4 for ND1, ND2, ND4, and ND5;

HKY+F+G4 for ND4L; TIM+F+I+G4 for CYTB; TIM2+F+I+G4 for COI, COII, COIII, and ND3; TN+F+I for ATP8; and TN+F+I+G4 for ND6.

Bayesian phylogenetic analyses were performed to estimate the divergence time using BEAST ver. 2.7.4. with the phylogenetic tree created using IQ-TREE as an initial tree. To constrain the age of the tree, three fossil calibration points were chosen:

1. For the divergence of the Vampyromorpha (*Vampyroteuthis infernalis*) and Octopoda, 162–250 million years ago (Ma) based on Strugnell et al. (2006).
2. For the divergence of the suborders Cirrata (*Opisthoteuthis massyae*) and Incirrata, 95 Ma based on Fuchs et al. (2009).
3. For the divergence of the genera *Tremoctopus* and *Argonauta*, 29 Ma based on the earliest record of the Argonautidae on Kobayashi (1954).

To estimate divergence times, I used an uncorrelated lognormal relaxed clock model incorporating the Calibrated Yule model (Heled & Drummond 2012) and a divergence rate of 0.02 substitutions per site per Myr (Pardo-Gandarillas et al. 2017). The substitution model most suitable for each region was selected according to BIC scores calculated by MEGA. The best fit models were selected separately for each protein-coding region as follows: TN93+G for ATP6, COII, ND3, and ND5; TN93+G for ND6; HKY+I for ATP8; HKY+G for ND4L; GTR+G+I for COI, COIII, ND1, ND2, and ND4; and GTR+G for

CYTB. MCMC runs were performed for 50 million steps sampled every 1,000 steps, with the first 5 million steps discarded as burn-in. These runs were repeated four times, and the log files were combined using LogCombiner. The convergence of the result of combined runs was verified by both visually inspecting the MCMC traces in Tracer ver. 1.6 to confirm the ESS > 200, and displayed using FigTree ver. 1.4.3.

4.3 Results

4.3.1 Population demography

4.3.1.1 Northwest Pacific

In the concatenated sequences, both of the neutral indices were negative and differed significantly from zero (Tajima's $D = -2.086$, Fu's $F_S = -15.217$). The mismatch distribution analysis showed a unimodal and recent population expansion (Table 9). The scores were small (0.008 in SSD and 0.121 in raggedness index), but the model of sudden expansion was rejected ($p < 0.05$). From the equation $t = \tau/(2\mu k)$, the expansion parameter τ was estimated to be 0.746 (95% CI: 0.512–1.166) and t = 0.033 Ma (95% CI: 0.023–0.052). The population expansion time was also validated by the BSP analysis using the concatenated sequences, which revealed that the population size expanded from about 0.01 Ma (Fig. 12A). The NeT was estimated to be 0.500.

Table 9. Summary of demographic indices for the Northwest Pacific (NWP) and Northeast Pacific (NEP)

Region	mtDNA gene	No. of samples	SSD		Harpending's raggedness index		tau (95% confidence interval)	t (95% confidence interval)
			Value	p	Value	p		
NWP	COI+ CYTB	259	0.008	0.029	0.121	0.010	0.746 (0.512–1.166)	0.033 (0.023–0.052)
NEP	COI	135	0.0003	0.019	0.971	0.959	0.300 (0.300–0.300)	0.029 (0.029–0.029)
	CYTB	116	0.017	0.019	0.140	0.0001	1.086 (0.793–1.598)	0.091 (0.066–0.134)

4.3.1.2 Northeast Pacific

The mismatch distribution analysis showed a recent unimodal population expansion (Fig. 12B, C). In COI, the scores were small (0.0003 in SSD and 0.971 in raggedness index), but only the raggedness index supported the model of sudden expansion ($p > 0.05$) (Table 9). The expansion parameter τ was estimated as to be 0.3 (95% CI: 0.3–0.3) and t = 0.029 Ma (95% CI: 0.029–0.029). In CYTB, the SSD (0.017) and raggedness index (0.140) were both small, but they rejected the model of sudden expansion ($p < 0.05$).

The expansion parameter τ was estimated to be 1.086 (95% CI: 0.793–1.598) and t = 0.091 Ma (95% CI: 0.066–0.134). The population expansion time was validated by BSP analysis using each sequence (COI and CYTB). In the COI analysis, the number of mutations was too small to estimate the demography (Fig. 12B). In the CYTB analysis, population size gradually expanded from about 0.01 Ma (Fig. 12C). The NeT was estimated to be 0.500.

4.3.2 Phylogenetic analysis

Using a divergence rate of 2% per million years, the genus *Enteroctopus* originated 16.22 Ma during the early Miocene (95% CI: 13.67–18.98 Ma) (Fig. 13). The branch of cryptic species appeared about 4.42 Ma during the Miocene (95% CI: 1.98–

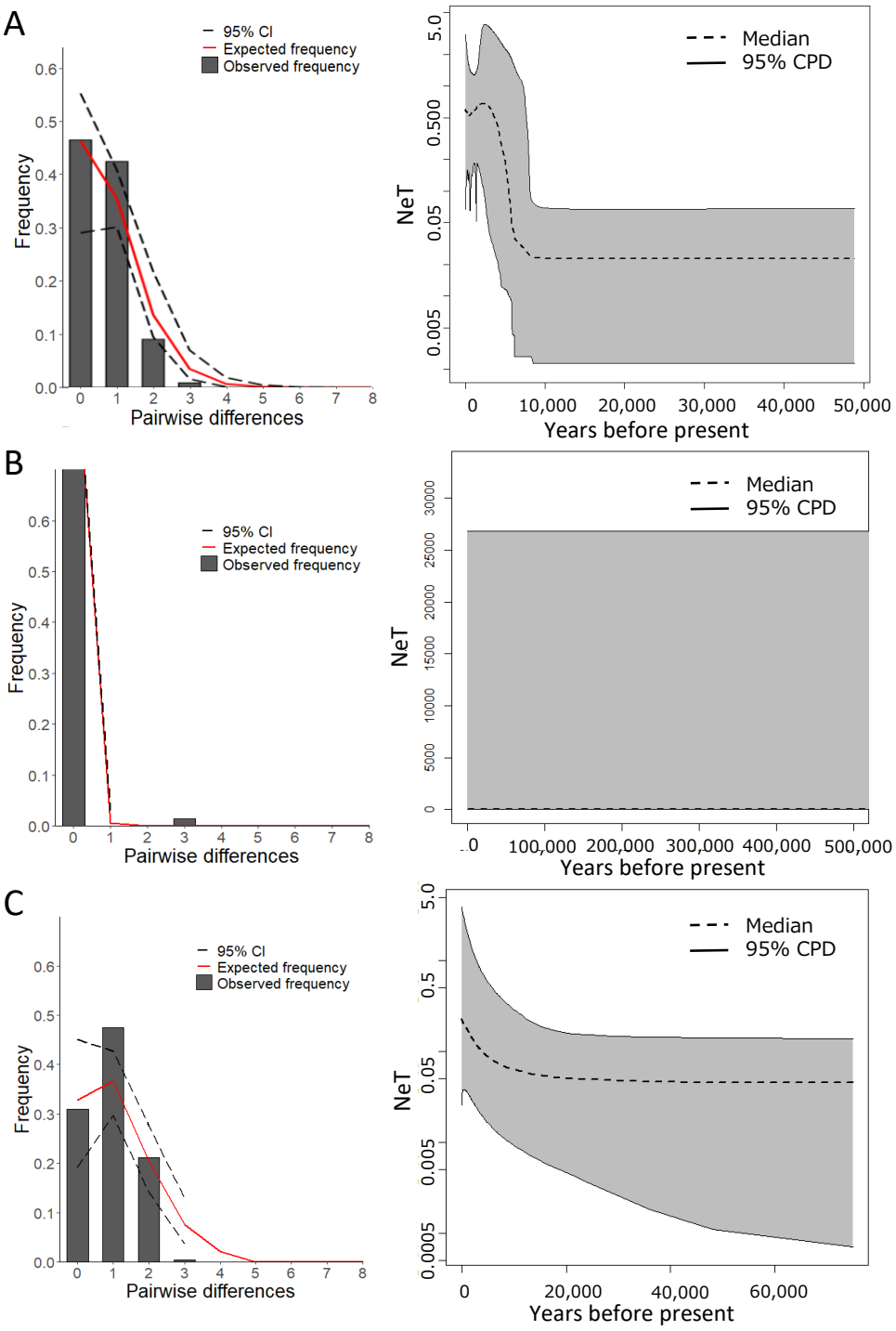


Figure 12. Demographic analyses of the NWP and NEP populations. Mismatch distribution diagrams under a sudden expansion model (left column) and Bayesian skyline plots showing change in effective population size (right column). CPD: conditional probability distribution. A; NWP with concatenate sequences (COI and CYTB: 1241 bp), B; NEP COI sequences (517 bp), C; NEP CYTB sequences (597 bp).

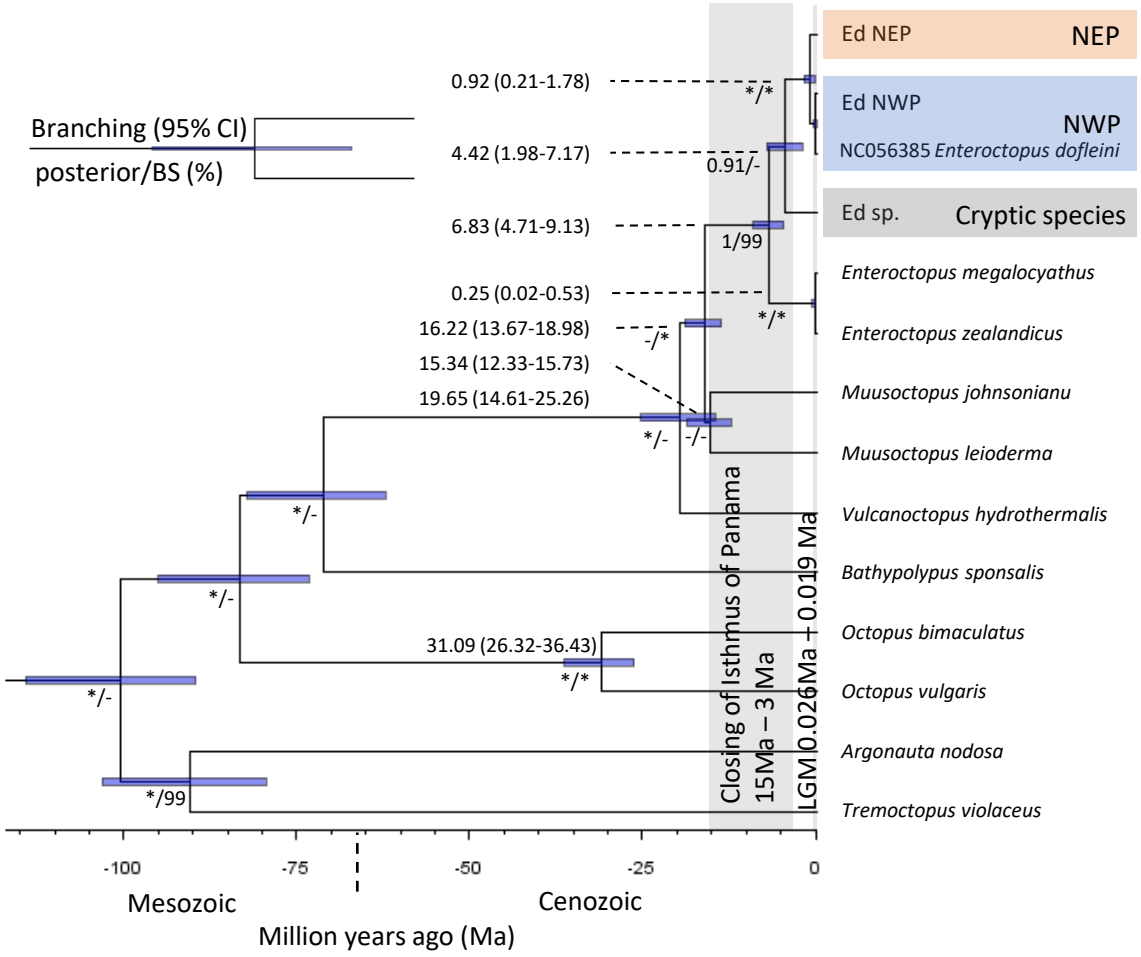


Figure 13. Phylogenetic time-calibrated tree of 14 cephalopods based on Bayesian inference of 13 protein coding regions from the mitochondrial gene with *Todarodes pacificus* as the outgroup. Divergence time estimates on X-axis are shown in millions of years. Colored bars on branch show 95% CI. The Bayesian posterior probabilities and bootstrap scores of the clades are shown below the nodes (posterior probability/ Bootstrap score). Only bootstrap values > 65 and posterior probabilities > 0.9 are shown. The bootstrap values = 100 and posterior probabilities = 1 are shown as *. The branching dates with 95% CI are shown over the nodes. Orange clade: Northeast Pacific. Blue clade: Northwest Pacific. Grey clade: cryptic species. The overall Bayesian phylogenetic tree appears in Appendix H.

7.17), and *E. dofleini* in the NEP and NWP branched about 0.92 Ma during the Pleistocene (95% CI: 0.21–1.78). The complete phylogenetic tree including the outgroup is presented in Appendix I.

4.4 Discussion

The results suggest that the divergence of the genus *Enteroctopus* occurred about 6.83 Ma, while the population expansion of *E. dofleini* occurred after the LGM in the NEP and NWP approximately 0.01 Ma. Previous studies on *Muusoctopus* used various evolutionary rates (Gleadall 2013; Ibañez et al. 2016; Lima et al. 2020b), but the results all indicated that *Enteroctopus* and *Muusoctopus* diverged from the common ancestor approximately 23 Ma, consistent with the results of this study (Fig. 13; 16.22 Ma [95% CI: 13.67 – 18.98]). Given that *Muusoctopus* is a deep-sea genus while *Enteroctopus* is distributed in relatively shallow depths around 200 meters, their divergence may have been influenced by inhabitable depth ranges and water temperature (Noro & Sakurai 2012; Sano & Bando 2015; Golikov et al. 2023). Like the change of temperature and sea level upwelling after the LGM, the environment factors were affected marine habitat connectivity, the availability of habitat space, and sediment sources (Kohfeld & Chase 2011; Pelto et al. 2018).

Ecological characteristics were also likely key factors in the divergence of these species, like the ability to maintain oxygen carrying capacity at sub-zero water temperatures by increasing blood hemocyanin levels as observed in an Antarctic species (*Pareledone charcoti*) (Oellermann et al. 2015). These adaptations may also be related to species diversity within each genus that shares similar distribution ranges within the family. Based on the WoRMS Editorial Board (2025), the genus *Muusoctopus* (25 species) is more diverse than the genus *Enteroctopus* (4 species). The divergence time also related to the diversity, and the divergence within the genus *Enteroctopus* may be more recent compared to that of *Muusoctopus*. Comparison of divergence time and ecological factors such as the fragmentation of habitat, adaptive traits, and larval morphology related to dispersal ability may be useful for understanding the pattern.

The formation of the Isthmus of Panama may have also played a role in the speciation of the *Enteroctopus* genus. The Isthmus of Panama is thought to have closed approximately 3 to 15 Ma (Bartoli et al. 2005; O'Dea et al. 2016), and species that diverged following this event are called transisthmian sister species pairs/complex (Marko 2002). *Muusoctopus* originated in the western Atlantic Ocean, and diverged during the Middle Miocene and Early Pliocene (mean range = 5–18 Ma) with the establishment of the isthmus, forming the diverse species seen today (Lima et al. 2020b). In the genus

Enteroctopus, *E. megalocyathus* and *E. zealandicus* diverged approximately 6.83 Ma (4.71–9.13 Ma), and a cryptic species of *E. dofleini* diverged around 4.42 Ma (1.98–7.17 Ma) in this study. However, the current distribution of the *Enteroctopus* genus is limited to the North Pacific, the southern tip of South America, New Zealand, and near the Isthmus of Panama (Gould 1852; Villanueva et al. 1991; Norman & Hochberg 2005). It is possible that *E. dofleini* spread widely across the North Pacific along the coast of North America following the closure of the Isthmus of Panama. In contrast, *E. megalocyathus* and *E. zealandicus* appear to have dispersed to the southern tip of South America, although the exact route remains unknown. While *E. magnificus* is reported to inhabit the western Atlantic, genetic information for this species is unavailable, so its phylogenetic status is unclear. Figure 14 summarizes the possible evolutionary history of the genus *Enteroctopus* based on the present study.

Cryptic species are genetically distinct but show small morphological differences (Bickford et al. 2007). In the NEP and NWP, a cryptic species and *E. dofleini* are thought to be sympatrically distributed (Toussaint et al. 2012a; Misawa et al. 2021). In this study the divergence of the cryptic species is estimated to have occurred approximately 4 Ma. In Barry et al. (2013) and Toussaint et al. (2012a), the cryptic species was collected mainly from areas with well-developed fjord topography, such as Prince William Sound, Stephen's

Passage. Individuals that entered these fjords following the LGM might have formed isolated populations with restricted connection where the cryptic species could remain. Cryptic species are often identified when speciation is recent and morphological variation is minimal (Hending 2025). Based on multiple microsatellite markers reported in Hollenbeck et al. (2017), there was no evidence of substantial genetic hybridization (n=21).

The results of the genetic analyses in this chapter suggest that the NEP and NWP populations both experienced recent population expansions about 0.01 Ma. Mismatch distribution analysis did not support the population expansion model except for the COI data in the NEP. However, the SSD and raggedness index values, which represent differences from the model, were below 0.01. In Jenkins et al. (2018), the netted dog whelk (*Nasarius reticulatus*) in northeast Atlantic showed small differences with an expansion model and insignificant differences, and they estimated that a recent expansion occurred (ca. 5 Ma). In other words, in species that have expanded relatively recently, population expansion signal may be weaker and thus more likely to show significant differences in comparison to an expansion model in the mismatch analysis.

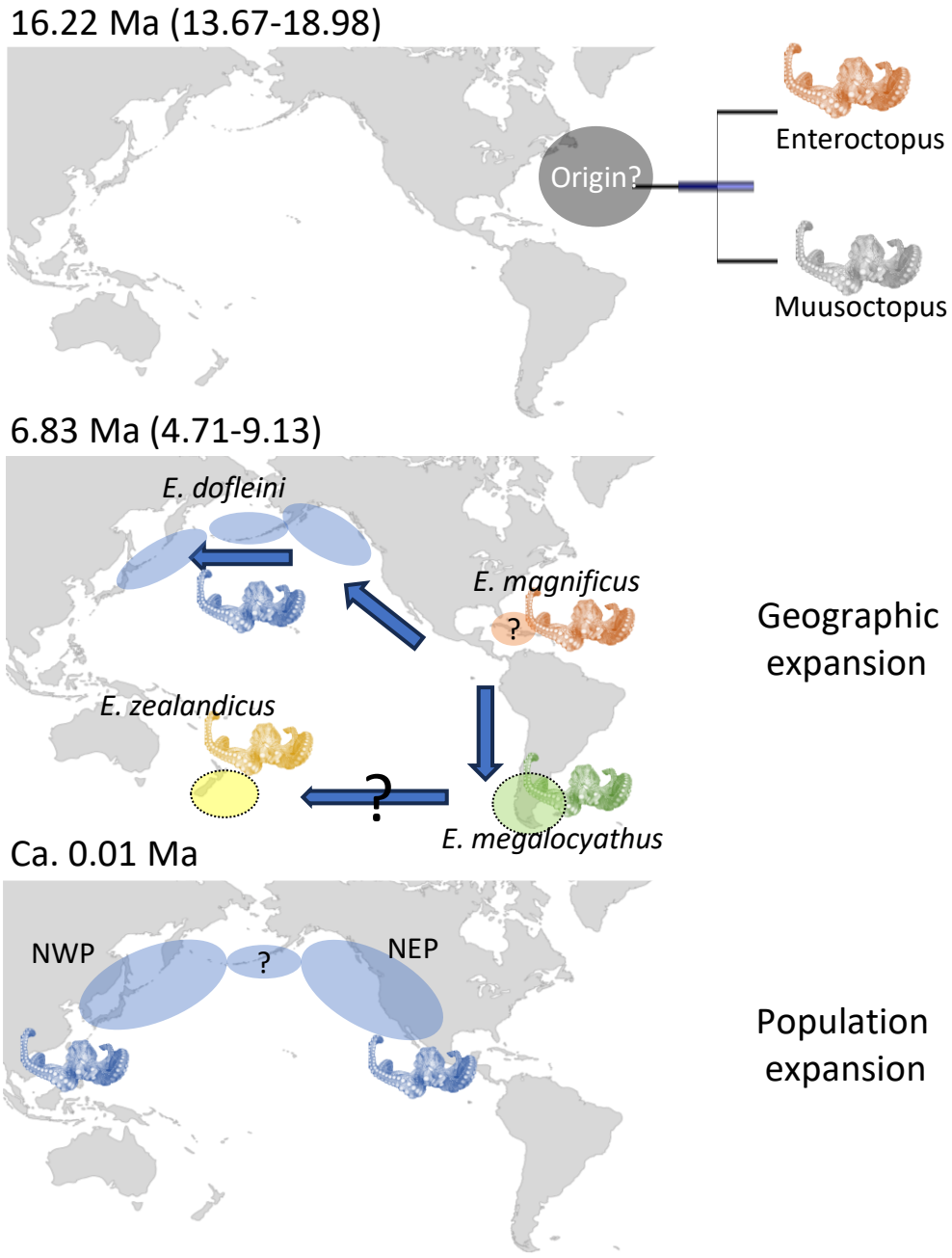


Figure 14. Schematic representation of proposed changes in historical distribution of *Enteroctopus dofleini* based on the results of this study. The estimated ages (in Ma) shown in the upper left of each panel are based on Fig. 13. First, around 16.22 Ma, lineages such as *Enteroctopus* and *Muusoctopus* diverged along the western Atlantic coast. The origin is assumed to be the western Atlantic, following Ibáñez et al. (2016). Subsequently, as the *Enteroctopus* genus expanded its range, speciation occurred, giving rise to the four extant species. Later, populations expanded in size along the Pacific Ocean, leading to its present distribution.

Chapter 5: General discussion

This dissertation was conducted to clarify the genetic diversity and population structure of the giant Pacific octopus (*Enteroctopus dofleini*) in the North Pacific, including waters around Hokkaido. In Chapter 2, analysis of two mitochondrial regions, COI and CYTB, revealed the genetic homogeneity of *E. dofleini* populations around Hokkaido, as well as their genetic differentiation from North American populations. In Chapter 3, neutral SNPs obtained through ddRAD sequencing and microsatellite analyses with seven markers were analyzed to further investigate the population structure around Japan. Similar to the mitochondrial analysis, no significant genetic differentiation between regions was found, and widespread gene flow among regions was estimated, which was presumably influenced by larval dispersal via ocean currents. Chapter 4 focused on demographic analysis, suggesting that populations expanded in both the NEP and NWP around 10,000 years ago. Additionally, phylogenetic analysis and divergence time estimation of the *Enteroctopus* genus were conducted, and I further explored the historical distribution of the genus and its relationship with the seaway closure of the Isthmus of Panama.

Understanding the population structure of a species plays a crucial role in resource management by helping monitor genetic diversity and maintaining effective population sizes. In Hokkaido, management of the fishery is based on the 11 geographic divisions proposed by Sano (2010), which were determined based in part on fluctuations in catch volume. Although populations corresponding to these divisions were not identified in this study, the results suggested the possibility of population-wide migration. Given previous studies that indicate post-settlement movement is relatively restricted (Noro & Sakurai 2012, Scheel & Bisson 2012, Bower et al. 2024), the broad gene flow within regions inferred from genetic analyses is likely maintained by larval dispersal or perhaps the occasional long-distance movement of a small number of subadults or adults.

While ecological and population differences between *E. dofleini* in the NEP and NWP had been previously suggested, sample numbers in past studies from Japan have been small, so comparisons between the regions have been asymmetrical (Toussaint et al. 2012a, Barry et al. 2013). In this study, a large dataset from Japan reaffirmed that although regional differences are subtle, they are clearly present. Furthermore, demographic analyses revealed for the first time that the differences between *E. dofleini* populations on opposite sides of the North Pacific may have arisen due to geographic distance and population expansion were occurred about 10,000 years ago.

Furthermore, while few studies have examined the phylogenetic relationships within the genus *Enteroctopus* (Pliego-Cárdenas et al. 2022), leaving its evolutionary history unclear, comparisons with previous studies of the closely related genus *Muusoctopus* (Gleadall 2013; Ibáñez et al. 2016, Lima et al. 2020b) suggest that events such as the closure of the Isthmus of Panama may have similarly influenced the dispersal and diversification of *Enteroctopus*.

A large number of samples from Japan were used in this study, but the sampling locations did not cover the entire distribution range of *E. dofleini* in Japan. Sampling in Hokkaido was concentrated in the north and south, with no sampling in the east, and while onboard sampling was conducted on the Pacific coast of Tohoku, only one sample was collected there (not included in this study). Since the cryptic species has been reported only from the coast of Tohoku in Japan (Misawa et al. 2021), further sampling efforts will be needed to better understand the distribution and population of this cryptic species. Due to the lack of samples from the Aleutian Islands and the Bering Sea, the division between the NEP and NWP remains speculative. Additionally, in the ddRAD-seq analysis, some regions had low sample numbers, which resulted in lower resolution in the analysis of the entire Japanese population. Furthermore, as sampling relied on specimens provided by

fishers, only weight and sex were recorded, and external morphological observations were not possible.

In terms of methodology, COI and CYTB, which include barcode regions that allow for interspecies differentiation, were selected as the mitochondrial analysis targets for comparison with previous studies and for phylogenetic estimation. However, due to their relatively low intraspecific diversity, these markers were not sufficient to assess variation within species. Given the high similarity between populations, using regions with higher diversity, such as control regions, may help detect potential populations by revealing intra-species differences. In ddRAD-seq, the choice of restriction enzymes is crucial. However, due to the limited genetic research that has been conducted on *E. dofleini*, I chose to use a fairly common combination of enzymes in this study. With the recent increase in genetic studies on cephalopods, selecting more appropriate enzymes could potentially enhance the information obtained.

E. dofleini has long been harvested as a stable resource, and recent efforts have been made to manage its population through fishing bans and weight restrictions to ensure its sustainable use. Despite these efforts, the resource levels of *E. dofleini* have been gradually declining in some areas. The resource reports produced by the Fisheries Research Agency (2022) suggest that, overall, the resource levels remain stable at medium

to high levels, but suggest resource levels have declined in southern areas of Hokkaido such as Tsugaru Strait and Okushiri Island (Hokkaido Research Organization 2019, Wakkanai Fisheries Research Institute 2022). In areas south of Hokkaido, including Aomori, Iwate, Miyagi, Fukushima, and Ibaraki prefectures, catch data indicate that Iwate and Miyagi, which are the main production areas on Honshu, have seen declining catches since 2016 (Fisheries Research and Education Agency 2021). While global warming and other factors may be contributing to this trend, the population genetics findings from this study could serve as valuable information for future resource management. Previous studies, such as Lima et al. (2020a), have shown that simulations can predict future changes in the distribution of *O. insularis*. Similarly, in the case of *E. dofleini*, larval dispersal, which supports genetic exchange between regions, is highly influenced by ocean currents. In addition to rising water temperatures, changes in dispersal routes may significantly impact future distribution. In the future, expanding the sampling effort, analyzing ocean currents, and incorporating ecological niche modeling with environmental factors (*e.g.*, Arjona et al. 2020, Boavida-Portugal et al. 2022) will allow for a more realistic examination of gene flow and population connectivity across space, taking into account the potential changes in distribution suggested in previous studies.

In conclusion, this study is the first to reveal the genetic homogeneity of *E. dofleini* populations around Hokkaido, suggesting extensive gene flow across a wide geographic range. It also indicates that speciation and population expansion events within the genus *Enteroctopus* may be associated with major historical events such as the formation of the Isthmus of Panama. Furthermore, the results of the study suggest a recent population expansion of *E. dofleini* approximately 10,000 years ago. These results highlight the need for broad and continuous monitoring for effective resource management of *E. dofleini*, to provide a fundamental basis for understanding its future population dynamics.

References

Aguirre N C, Filippi C V, Vera P A, Puebla A F, Zaina G, Lia V V, Marcucci Poltri S N, Paniego N B (2023) Double digest restriction-site associated DNA sequencing (ddRADseq) technology. In: *Plant Genotyping: Methods and Protocols*, Shavrukov Y (eds), vol. 2638, pp. 37–57. New York, NY: Springer US.

Amor M D, Norman M D, Roura A, Leite T S, Gleadall I G, Reid A, Perales-Raya C, Lu C C, Silvey C J, Vidal E A G, Hochberg F G, Zheng X, Strugnell J M (2017) Morphological assessment of the *Octopus vulgaris* species complex evaluated in light of molecular-based phylogenetic inferences. *Zoologica Scripta*, 46, 275–288.

Andrews K R, Good J M, Miller M R, Luikart G, Hohenlohe P A (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, 17, 81–92.

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Arjona Y, Fernández-López J, Navascués M, Alvarez N, Nogales M, Vargas P (2020) Linking seascape with landscape genetics: Oceanic currents favour colonization across the Galápagos Islands by a coastal plant. *Journal of Biogeography*, 47, 2622–2633.

Avise J C (2000) *Phylogeography: the history and formation of species*. Harvard University Press.

Azuma N, Chiba S (2017) Genetic population structure of the Hokkai shrimp *Pandalus latirostris* Rathbun, 1902 (Decapoda: Caridea: Pandalidae) from Zostera seagrass beds in Hokkaido, Japan. *Journal of Crustacean Biology*, 38, 147–155.

Azuma N, Kunihiro Y, Sasaki J, Mihara E, Mihara Y, Yasunaga T, Jin D-H, Abe S (2008) Genetic variation and population structure of hair crab (*Erimacrus isenbeckii*) in Japan inferred from mitochondrial DNA sequence analysis. *Marine Biotechnology*, 10, 39–48.

Azuma N, Zaslavskaya N I, Yamazaki T, Nobetsu T, Chiba S (2017) Phylogeography of *Littorina sitkana* in the northwestern Pacific Ocean: evidence of eastward trans-Pacific colonization after the last glacial maximum. *Genetica*, 145, 139–149.

Baird N A, Etter P D, Atwood T S, Currey M C, Shiver A L, Lewis Z A, Selker E U, Cresko W A, Johnson E A (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3: e3376.

Barry P D, Tamone S L, Tallmon D A (2013) A complex pattern of population structure in the North Pacific giant octopus *Enteroctopus dofleini* (Wülker, 1910). *Journal of Molluscan Studies,* 79, 133–138.

Bartoli G, Sarnthein M, Weinelt M, Erlenkeuser H, Garbe-Schönberg D, Lea D W (2005) Final closure of Panama and the onset of northern hemisphere glaciation. *Earth and Planetary Science Letters*, 237, 33–44.

Begg G A, Friedland K D, Pearce J B (1999) Stock identification and its role in stock assessment and fisheries management: an overview. *Fisheries Research*, 43, 1–8.

Bein, B, Lima F D, Lazzarotto H, Rocha L A, Leite T S, Lima S M Q, Pereira R J (2023) Population genomics of an Octopus species identify oceanographic barriers and inbreeding patterns. *Marine Biology*, 170, 161.

Benham W B (1944) The octopodous mollusca of New Zealand 4. *Transactions of the Royal Society of New Zealand*, 73, 255–261.

Bickford D, Lohman D J, Sodhi N S, Ng P K L, Meier R, Winker K, Ingram K K, Das I (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, 22, 148–155.

Boavida-Portugal J, Guilhaumon F, Rosa R, Araújo M B (2022) Global patterns of coastal cephalopod diversity under climate change. *Frontiers in Marine Science*, 8, 740781.

Bohonak A J (1999) Dispersal, gene flow, and population structure. *The Quarterly Review of Biology*, 74, 21–45.

Boletzky S V (2003) Biology of early life stages in cephalopod molluscs. *Advances in Marine Biology*, 44, 143–203.

Bower J R, Noro K, Sakurai Y, Scheel D (2024) *Enteroctopus dofleini*, Giant Pacific octopus. In: *Octopus Biology and Ecology*, Rosa R, Gleadall I G, Pierce G J, Villanueva R (eds.), pp. 369–395. San Diego, Academic Press.

Canino M F, Spies I B, Cunningham K M, Hauser L, Grant W S (2010) Multiple ice-age refugia in Pacific cod, *Gadus macrocephalus*. *Molecular Ecology*, 19, 4339–4351.

Cassone B J, Boulding E G (2006) Genetic structure and phylogeography of the lined shore crab, *Pachygrapsus crassipes*, along the northeastern and western Pacific coasts. *Marine Biology*, 149, 213–226.

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17, 540-552.

Catchen J, Hohenlohe P A, Bassham S, Amores A, Cresko W A (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140.

Chang C C, Chow C C, Tellier L C, Vattikuti S, Purcell S M, Lee J J (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4, s13742–015–0047–8.

Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor, *Bioinformatics*, 34, i884–i890.

Cheng S H, Gold M, Rodriguez N, Barber P H (2021) Genome-wide SNPs reveal complex fine scale population structure in the California market squid fishery (*Doryteuthis opalescens*). *Conservation Genetics*, 22, 97–110.

Chust G, Villarino E, Chenuil A, Irigoien X, Bizsel N, Bode A, Broms C, Claus S, Fernández de Puelles M L, Fonda-Umani S, Hoarau G, Mazzocchi G M, Mozetič P, Vandepitte L, Veríssimo H, Zervoudaki S, Borja A (2016) Dispersal similarly shapes both population genetics and community patterns in the marine realm. *Scientific Reports*, 6, 28730.

Clark P U, Mix A C (2002) Ice sheets and sea level of the last glacial maximum. *Quaternary Science Reviews*, 21, 1–7.

Conrath C L, Conners M E (2014) Aspects of the reproductive biology of the North Pacific giant octopus (*Enteroctopus dofleini*) in the Gulf of Alaska. *Fishery Bulletin*, 112, 253–260.

Cowen R K, Sponaugle S (2009) Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, 1, 443–466.

Cox L N, Zaslavskaya N I, Marko P B (2014) Phylogeography and trans-Pacific divergence of the rocky shore gastropod *Nucella lima*. *Journal of Biogeography,* 41, 615–627.

Dan S, Iwasaki H, Takasugi A, Yamazaki H, Hamasaki K (2018) An upwelling system for culturing common octopus paralarvae and its combined effect with supplying natural zooplankton on paralarval survival and growth. *Aquaculture*, 495, 98–105.

Davey J W, Blaxter M L (2010) RAD-Seq: next-generation population genetics. *Briefings in Functional Genomics*, 9.5-6, 416–423.

Drummond A J, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7, 214.

Duforet-Frebourg N, Bazin E, Blum M G B (2014) Genome scans for detecting footprints of local adaptation using a Bayesian factor model. *Molecular Biology and Evolution*, 31, 2483–2495.

Earl D A, von Holdt B M (2011) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361.

Esser E A, Pringle J M, Byers J E (2023) Neither larval duration nor dispersal distance predict spatial genetic diversity in planktonic dispersing species. *Marine Ecology Progress Series*, 721, 161–167.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology,* 14, 2611–2620.

Excoffier L, Lischer H E L (2010) Arlequin suite ver. 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564‑567.

FAO FishStat (2025) Global Capture Production. In: Fisheries and Aquaculture. [Accessed January 13, 2025]. https://www.fao.org/fishery/en/collection/capture?lang=en

Fernández-Álvarez F Á, Sánchez P, Villanueva R (2021) Morphological and molecular assessments of bobtail squids (Cephalopoda: Sepiolidae) reveal a hidden history of biodiversity. *Frontiers in Marine Science*, 7, 632261.

Fischer M C, Rellstab C, Leuzinger M, Roumet M, Gugerli F, Shimizu K K, Holderegger R, Widmer A (2017) Estimating genomic diversity and population differentiation–an empirical comparison of microsatellite and SNP variation in *Arabidopsis halleri*. *BMC Genomics*, 18, 1-15.

Fisheries Research and Education Agency (2021) Resource assessment report of giant Pacific octopus in the waters around northern part of Japan. *Fisheries Research and Education Agency*. (in Japanese)

Fisheries Research and Education Agency (2022) Resource assessment report in the waters around northern part of Japan, Hokkaido. *Fisheries Research and Education Agency*. (in Japanese)

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3, 294–299.

Fu Y-X (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, 147, 915–925.

Fuchs D, Bracchi G, Weis R (2009) New octopods (Cephalopoda: Coleoidea) from the late Cretaceous (Upper Cenomanian) of Hâkel and Hâdjoula, Lebanon. *Palaeontology*, 52, 65–81.

García-Mayoral E, Silva C N, Ramilo A, Roura Á, Moreno A, Strugnell J M, González Á F (2024) Population connectivity of the European squid *Loligo vulgaris* along the West Iberian Peninsula coast: comparing mtDNA and SNPs. *Marine Biology*, *171*, 237.

Gleadall I G (2013) A molecular sequence proxy for *Muusoctopus januarii* and calibration of recent divergence among a group of mesobenthic octopuses. *Journal of Experimental Marine Biology and Ecology*, 447, 106–122.

Golikov A V, Gudmundsson G, Blicher M E, Jørgensen L L, Korneeva E I, Olafsdottir S H, Shagimardanova E I, Shigapova L H, Zakharov D V, Zimina O L, Sabirov R M (2023) A review of the genus *Muusoctopus* (Cephalopoda: Octopoda) from Arctic waters. *Zoological Letters*, 9, 21.

Gould A A (1852) *Mollusca and shells, with an atlas of plates. United States Exploring Expedition: during the years 1838, 1839, 1840, 1841, 1842. Under the command of C. Wilkes, U.S.N. Vol. XII*. Sherman, Philadelphia. xv + 510 pp., atlas 16 pp. + 52 pl.

Grant W S, Cheng W (2012) Incorporating deep and shallow components of genetic structure into the management of Alaskan red king crab. *Evolutionary Applications*, 5, 820–837.

Griffiths S M, Butler M J, Behringer D C, Pérez T, Preziosi R F (2021) Oceanographic features and limited dispersal shape the population genetic structure of the vase sponge *Ircinia campana* in the Greater Caribbean. *Heredity,* 126, 63–76.

Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepoittevin T, Malausa T, Revardel E, Salin F, Petit R J (2011) Current trends in microsatellite genotyping. *Molecular Ecology Resources*, 11, 591-611.

Harpending H C, Sherry S T, Rogers A R, Stoneking M (1993) Genetic structure of ancient human populations. *Current Anthropology,* 34, 483–496.

Heled J, Drummond A J (2012) Calibrated tree priors for relaxed phylogenetics and divergence time estimation. *Systematic Biology,* 61, 138–49.

Hellberg M E (2009) Gene flow and isolation among populations of marine animals. *Annual Review of Ecology, Evolution and Systematics*, 40, 291–310.

Hending D (2025) Cryptic species conservation: a review. *Biological Reviews*, 100, 258–274.

Hochberg F G (1998) Class Cephalopoda. In: *Taxonomic atlas of the benthic fauna of the Santa Maria basin and the western Santa Barbara channel: The Mollusca, Part 1. Santa Barbara Museum of Natural History*, Scott P V, Blake J A (eds.), vol. 8, pp. 175–236. Santa Barbara, CA.

Hokkaido Research Organization (2019) Resource assessment report of giant Pacific octopus in the waters around Hokkaido. *Hokkaido Research Organization Fisheries Research Department.* pp. 41.1–41.17. (in Japanese)

Hollenbeck N, Scheel D (2017) Body patterns of the frilled giant Pacific octopus, a new species of octopus from Prince William Sound, AK. *American Malacological Bulletin,* 35, 134–144.

Hollenbeck N, Scheel D, Gravley M C, Sage G K, Toussaint R, Talbot S L (2017) Use of swabs for sampling epithelial cells for molecular genetics analyses in *Enteroctopus*. *American Malacological Bulletin,* 35, 145–157.

Huson D H, Bryant, D (2006) Application of phylogenetic networks in evolutionary studies, *Molecular Biology and Evolution*, 23, 254–267.

Ibáñez C M, Pardo-Gandarillas M C, Peña F, GleadalI I G, Poulin E, Sellanes J (2016) Phylogeny and biogeography of *Muusoctopus* (Cephalopoda: Enteroctopodidae). *Zoologica Scripta*, 45, 494–503.

Jenkins T L, Castilho R, Stevens J R (2018) Meta-analysis of northeast Atlantic marine taxa shows contrasting phylogeographic patterns following post-LGM expansions. *PeerJ* 6: e5684.

Joh M, Miyoshi K, Sato M, Sano M (2017) Seasonal migration, growth, and exposed water temperature of immature giant octopus *Enteroctopus dofleini* using mark and recapture experiment and seasonal change in bottom water temperature in the fishing ground. *Bulletin of the Japanese Society of Fisheries Oceanography (Japan),* 81, 50–59. (in Japanese with English abstract)

Jombart T, Ahmed I (2011) *adegenet 1.3-1*: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27, 3070–3071.

Kanamaru S, Yamashita Y (1969) The fishery biology for the octopus, "mizu-dako", (*Paraoctopus hongkongensis* (Hoyle)). (I) Summer movements in Onishika area of North-western part of Hokkaido. *Bulletin of the Hokkaido Regional Fisheries Research,* 35, 178–197. (in Japanese with English abstract)

Kaneko N, Kubodera T, Iguchi A (2011) Taxonomic study of shallow-water octopuses (Cephalopoda: Octopodidae) in Japan and adjacent waters using mitochondrial genes with perspectives on octopus DNA barcoding. *Malacologia*, 54, 97–108.

Kawamura K, Miyake T, Obata M, Aoki H, Komaru A (2017) Population demography and genetic characteristics of the Pacific oyster *Crassostrea gigas* in Japan. *Biochemical Systematics and Ecology*, 70, 211–221.

Kobayashi G, Itoh H, Kanaya G, Abe H, Kojima S (2023) Genetically isolated population of the coastal species with high dispersal potential: the case of the sentinel crab *Macrophthalmus japonicus* (Brachyura: Macrophthalmidae) in Japan. *Plankton and Benthos Research*, 18, 13–20.

Kobayashi T (1954) A new palaeogene paracenoceratoid from southern Kyushu in Japan. *Japanese Journal of Geology and Geography,* 24, 181–184.

Kohfeld K E, Chase Z (2011) Controls on deglacial changes in biogenic fluxes in the North Pacific Ocean. *Quaternary Science Reviews*, 30, 3350–3363.

Kojima S, Kojima S, Sakuma K, Yanagimoto T (2014) Contrasting genetic population structures between congeneric flounder species, *Hippoglossoides dubius* and *H. pinetorum*. *Plankton and Benthos Research*, 9, 99–104.

Kopelman N M, Mayzel J, Jakobsson M, Rosenberg N A, Mayrose I (2015) Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15, 1179–1191.

Kubodera T (1991) Distribution and abundance of the early life stages of octopus, *Octopus dofleini* Wülker, 1910 in the North Pacific. *Bulletin of Marine Science,* 49, 235–244.

Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, 33, 1870–1874.

Larson S, Ramsay C, Cosgrove J A (2015) Multiple paternity and preliminary population genetics of giant Pacific octopuses, *Enteroctopus dofleini*, in Oregon, Washington and the southeast coast of Vancouver Island, BC. *Diversity*, 7, 195–205.

Larson W A, Seeb L W, Everett M V, Waples R K, Templin W D, Seeb J E (2014) Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evolutionary Applications*, 7, 355–369.

Leigh J W, Bryant D (2015) POPART: full-feature software for haplotype network construction. *Methods in Ecology & Evolution*, 6, 1110–1116.

Leite T S, Haimovici M, Molina W, Warnke K (2008) Morphological and genetic description of *Octopus insularis*, a new cryptic species in the *Octopus vulgaris* complex (Cephalopoda: Octopodidae) from the tropical southwestern Atlantic. *Journal of Molluscan Studies*, 74, 63–74.

Lima F D, Ángeles-González L E, Leite T S, Lima S M Q (2020a) Global climate changes over time shape the environmental niche distribution of *Octopus insularis* in the Atlantic Ocean. *Marine Ecology Progress Series*, 652, 111–121.

Lima F D, Strugnell J M, Leite T S, Lima S M (2020b) A biogeographic framework of octopod species diversification: the role of the isthmus of Panama. *PeerJ*, 8, e8691.

Liu M, Lin L, Gao T, Yanagimoto T, Sakurai Y, Grant W S (2012) What maintains the central North Pacific genetic discontinuity in Pacific herring? *PLoS One*, 7, e50340.

Madeira A G, Tsuda Y, Nagano Y, Iwasaki T, Zucchi M I, Kajita T, Mori G M (2023) The role of oceanic currents in the dispersal and connectivity of the mangrove Rhizophora mangle on the Southwest Atlantic region. *Molecular Ecology Resources*, 1–16.

Mann D H, Gaglioti B V (2024) The northeast Pacific Ocean and northwest coast of North America within the global climate system, 29,000 to 11,700 years ago. *Earth-Science Reviews*, 254, 104782.

Marko P B (2002) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the isthmus of Panama. *Molecular Biology and Evolution*, 19, 2005–2021.

Marko P B, Hoffman J M, Emme S A, McGovern T M, Keever C C, Cox L N (2010) The "Expansion-Contraction" model of Pleistocene biogeography: Rocky shores suffer a sea change? *Molecular Ecology*, 19, 146–169.

Melis R, Vacca L, Cuccu D, Mereu M, Cau A, Follesa M C, Cannas R (2018) Genetic population structure and phylogeny of the common octopus *Octopus vulgaris* Cuvier, 1797 in the western Mediterranean Sea through nuclear and mitochondrial markers. *Hydrobiologia*, 807, 277–296.

Minh B Q, Schmidt H A, Chernomor O, Schrempf D, Woodhams M D, von Haeseler A, Lanfear R (2020) IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biological Evolution*, 37, 1530–1534.

Misawa R, Narimatsu Y, Suzuki Y, Morikawa E, Tokioka S, Nagao J, Saitoh K (2021) DNA barcoding of the northeast Pacific Ocean octopus family - example in the North area -. *Northeast Bottom Fish Research*, 41, 9–15. (in Japanese)

Morgan J A, Dunning M C, Ovenden J R, Broderick D, Street R, Davies J, Hall K C (2024) Molecular delimitation of cryptic Australian squid species of the genus Uroteuthis Rehder, 1945 (Cephalopoda: Loliginidae), provides a baseline of diversity to resolve classification challenges throughout the Indo-Pacific. *Journal of Molluscan Studies*, 90, eyae028.

Morse P, Kjeldsen S R, Meekan M G, Mccormick M I, Finn J K, Huffard C L, Zenger K R (2018) Genome-wide comparisons reveal a clinal species pattern within a holobenthic octopod—the Australian Southern blue-ringed octopus, *Hapalochlaena maculosa* (Cephalopoda: Octopodidae). *Ecology and Evolution*, 8, 2253–2267.

Mussmann M, Douglas M R, Chafin T K, Douglas M E (2019) BA3-SNPs: Contemporary migration reconfigured in BayesAss for next-generation sequence data. *Methods in Ecology and Evolution*, 10, 1808–1813.

Norman M D, Hochberg F G (2005) The current state of octopus taxonomy. *Phuket Marine Biological Center Research Bulletin,* 66, 127–154.

Norman M, Finn J, Hochberg F G (2016) Family Octopodidae. In: *FAO species catalogue for fishery purposes. No. 4*: *Cephalopods of the world. An annotated and illustrated*

catalogue of cephalopod species known to date, Jareb P, Roper C F E, Norman M D, Finn J (eds.), vol. 3, pp. 36–215. Rome, FAO.

Noro K, Sakurai Y (2012) Migration, distribution and growth patterns of the North Pacific giant octopus *Enteroctopus dofleini* in waters adjacent to Tsugaru Strait. *Aquaculture Science*, 60, 429–443. (in Japanese with English abstract)

O'Dea A, Lessios H A, Coates A G, Eytan R I, Restrepo-Moreno S A, Cione A L, Collins L S, de Queiroz A, Farris D W, Norris R D, Stallard R F, Woodburne M O, Aguilera O, Aubry M P, Berggren W A, Budd A F, Cozzuol M A, Coppard S E, Duque-Caro H, Finnegan S, Gasparini G M, Grossman E L, Johnson K G, Keigwin L D, Knowlton N, Leigh E G, Leonard-Pingel J S, Marko P B, Pyenson N D, Rachello-Dolmen P G, Soibelzon E, Soibelzon L, Todd J A, Vermeij G J, Jackson J B (2016) Formation of the Isthmus of Panama. *Science Advances*, 2, e1600883.

Oellermann M, Lieb B, Pörtner H O, Semmens J M, Mark F C (2015) Blue blood on ice: modulated blood oxygen transport facilitates cold compensation and eurythermy in an Antarctic octopod. *Frontiers in Zoology*, 12, 1–17.

Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Borman T, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2025) *vegan: Community Ecology Package*. R package version 2.7-0.

Okubo S (1980) Culture of the paralarvae and juveniles of *Octopus dofleini* in an aquarium for one year and two months. *Shima Marineland Quarterly*, 25, 4–5. (in Japanese)

Okubo S (1989) *Octopus dofleini*. Breeding of saltwater fish. In: *Kaisuigyo no hanshoku: Sodatete miyou umi no ikimono* [Marine fish reproduction: Let's raise, creatures of the sea], Suzuki K, Komatsu S (eds.), pp. 232–234. Japan, Midori Shobo. (in Japanese)

Orozco-Ruiz A M, Galván-Tirado C, Orlov A M, Orlova S Y, García-De León F J (2023) Genetic analyses reveal a non-panmictic genetic structure in the sablefish *Anoplopoma fimbria* in the northern Pacific. *ICES Journal of Marine Science*, 80, 1319–1328.

Palsbøll P J, Berube M, Allendorf F W (2007) Identification of management units using population genetic data. *Trends in Ecology and Evolution*, 22, 11-16.

Palumbi S R (1992) Marine speciation on a small planet. *Trends in Ecology & Evolution,* 7, 114–118.

Palumbi S R (1994) Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, 25, 547–572.

Pardo-Gandarillas M C, Ibáñez C M, Yamashiro C, Méndez M A, Poulin E (2017) Demographic inference and genetic diversity of *Octopus mimus* (Cephalopoda: Octopodidae) throughout the Humboldt current system. *Hydrobiologia*, 808, 125–135.

Peakall R, Smouse P E (2012) GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics,* 28, 2537–2539.

Pelto B M, Caissie B E, Petsch S T, Brigham-Grette J (2018) Oceanographic and climatic change in the Bering Sea, last glacial maximum to Holocene. *Paleoceanography and Paleoclimatology*, 33, 93–111.

Pérez-Alvare M J, Olavarría C, Moraga R, Baker C S, Hamner R M, Poulin E (2016) Historical dimensions of population structure in a continuously distributed marine species: The case of the endemic Chilean dolphin. *Scientific Reports*, 6, 35507.

Peterson B K, Weber J N, Kay E H, Fisher H S, Hoekstra H E (2012) Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7, e37135.

Pickford G E (1964) *Octopus dofleini* (Wülker), the giant octopus of the North Pacific. *Bulletin of the Bingham Oceanographic Collection*, 19, 1–70.

Pliego-Cárdenas R, Schofield-Astorga D C, Acuña-Gómez E P, de los Angeles Barriga-Sosa I (2022) A first look at the genetic diversity of *Enteroctopus megalocyathus* (Cephalopoda: Enteroctopodidae) captured by the king crab fishery in the south of Chile. *Journal of the Marine Biological Association of the United Kingdom*, 102, 377–385.

Pritchard J K, Stephens M, Rosenberg N A, Donnelly P (2000) Association mapping in structured populations. *The American Journal of Human Genetics*, 67, 170–181

Privé F, Luu K, Vilhjálmsson B J, Blum M G (2020) Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. *Molecular Biology and Evolution*, 37, 2153–2154.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M A, Bender D, Maller J, Sklar P, Bakker P I W, Daly M J, Sham P C (2007) PLINK: A tool set for whole-genome

association and population-based linkage analyses. *The American Journal of Human Genetics*, 81, 559–575.

R Core Team (2024) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Rae J W, Gray W R, Wills R C J, Eisenman I, Fitzhugh B, Fotheringham M, Littley E F M, Rafter P A, Rees-Owen R, Ridgwell A, Taylor B, Burke A (2020) Overturning circulation, nutrient limitation, and warming in the Glacial North Pacific. *Science Advances*, 6, eabd1654.

Rambaut A, Suchard M, Xie D, Drummond A J (2014) Tracer v.1.6. Computer program and documentation distributed by the author. http://beast.bio.ed.ac.uk/Tracer [Accessed June 7, 2024]

Rigby P R, Sakurai Y (2005) Multidimensional tracking of giant Pacific octopuses in northern Japan reveals unexpected foraging behaviour. *Marine Technology Society Journal*, 39, 64-67.

Ronquist F, Teslenko M, van der Mark P, Ayres D L, Darling A, Höhna S, Larget B, Liu L, Suchard M A, Huelsenbeck J P (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61, 539–542.

Roura A, Amor M, González A F, Guerra Á, Barton E D, Strugnell J M (2019) Oceanographic processes shape genetic signatures of planktonic cephalopod paralarvae in two upwelling regions. *Progress in Oceanography*, 170, 11–27.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, 145, 1219–1228.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio J C, Guirao-Rico S, Librado P, Ramos-Onsins S E, Sánchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, 34, 3299–3302.

Sales J B D L, Rodrigues-Filho L F D S, Ferreira Y D S, Carneiro J, Asp N E, Shaw P W, Haimovici M, Markaida U, Ready J, Schneider H, Sampaio I (2017) Divergence of cryptic species of *Doryteuthis plei* Blainville, 1823 (Loliginidae, Cephalopoda) in the Western Atlantic Ocean is associated with the formation of the Caribbean Sea. *Molecular Phylogenetics and Evolution*, 106, 44–54.

Sano M (2010) Geographical distribution of *Octopus dofleini* fisheries assessed for fisheries management using a geographic information system in the coastal areas around Hokkaido. *Scientific Reports of Hokkaido Fisheries Experimental Station* 77, 73–82. (in Japanese with English abstract)

Sano M, Bando T (2015) Seasonal migration of North Pacific giant octopus *Enteroctopus dofleini* in the Soya/La Pérouse Strait. *Nippon Suisan Gakkaishi (Japan)*, 81, 27–42. (in Japanese with English abstract)

Sato K, Yorita T (1999) Fishing trends and migration of the North Pacific giant octopus *Octopus dofleini* in Tsugaru Straits. *Scientific Reports of Hokkaido Fisheries Experimental Station*, 56, 119–124 (in Japanese with English abstract)

Sauer, W H H, Gleadall I G, Downey-Breedt N, Doubleday Z, Gillespie G, Haimovici M, Ibáñez C M, Katugin O N, Leporati S, Lipinski M R, Markaida U, Ramos J E, Rosa R, Villanueva R, Arguelles J, Briceño F A, Carrasco S A, Che L J, Chen C, Cisneros R, Conners E, Crespi-Abril A C, Kulik V V, Drobyazin E N, Emery T, Fernández-Álvarez F A, Furuya H, González L W, Gough C, Krishnan P, Kumar B, Leite T, Lu C, Mohamed K S, Nabhitabhanta J, Noro K, Petchkamnerd J, Putra D, Rocliffe S, Sajikumar K K, Sakaguchi H, Samuel D, Sasikumar G, Wada T, Zheng X, Tian Y, Pang Y, Yamrungrueng A, Pecl G (2019). World octopus fisheries. *Reviews in Fisheries Science & Aquaculture*. 29, 279–429.

Scheel D, Bisson L (2012) Movement patterns of giant Pacific octopuses, *Enteroctopus dofleini* (Wülker, 1910). *Journal of Experimental Marine Biology and Ecology,* 416, 21–31.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, 393–430.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139, 457-462.

Sotil G, Cueva D (2019) RAD sequencing technology for the evaluation of the population genome variability of the jumbo flying squid *Dosidicus gigas* in the Peruvian jurisdictional waters. Working paper presented to the 7th meeting of the Scientific Committee of the SPRFMO, La Havana, Cuba, 7–12 October 2019. SPRFMO Document SC7-SQxx (TBA): 7 p

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312–1313.

Strugnell J, Jackson J, Drummond A J, Cooper A (2006) Divergence time estimates for major cephalopod groups: evidence from multiple genes. *Cladistics*, 22, 89–96

Strugnell J M, Norman M D, Vecchione M, Guzik M, Allcock A L (2014) The ink sac clouds octopod evolutionary history. *Hydrobiologia,* 725, 215–235.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123, 585–595.

Takeuchi T, Masaoka T, Aoki H, Koyanagi R, Fujie M, Satoh N (2020) Divergent northern and southern populations and demographic history of the pearl oyster in the western Pacific revealed with genomic SNPs. *Evolutionary Applications*, 13, 837–853.

Takumiya M, Kobayashi M, Tsuneki K, Furuya H (2005) Phylogenetic relationships among major species of Japanese coleoid cephalopods (Mollusca: Cephalopoda) using three mitochondrial DNA sequences. *Zoological Science*, 22, 147–155.

Tanabe A S (2007) Kakusan: a computer program to automate the selection of a nucleotide substitution model and the configuration of a mixed model on multilocus data. *Molecular Ecology Notes*, 7, 962–964.

Tanaka S, Takahashi K (2005) Late Quaternary paleoceanographic changes in the Bering Sea and the western subarctic Pacific based on radiolarian assemblages. *Deep Sea Research Part II: Topical Studies in Oceanography*, 52, 2131–2149.

Theissinger K, Fernandes C, Formenti G, Bista I, Berg P R, Bleidorn C, Bombarely A, Crottini A, Gallo G R, Godoy J A, Jentoft S, Malukiewicz J, Mouton A, Oomen R A, Paez S, Palsbøll P J, Pampoulie C, Ruiz-López M J, Secomandi S, Svardal H, Theofanopoulou C, de Vries J, Waldvogel A M, Zhang G, Jarvis E D, Bálint M, Ciofi C, Waterhouse R M, Mazzoni C J, Höglund J (2023) European Reference Genome Atlas Consortium. How genomics can help biodiversity conservation. *Trends in Genetics*, 39, 545–559.

Timm L E, Bracken-Grissom H D, Sosnowski A, Breitbart M, Vecchione M, Judkins H (2020) Population genomics of three deep-sea cephalopod species reveals connectivity between the Gulf of Mexico and northwestern Atlantic Ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 158, 103222.

Toussaint R K, Scheel D, Sage G K, Talbot SL (2012a) Nuclear and mitochondrial markers reveal evidence for genetically segregated cryptic speciation in giant Pacific octopuses from Prince William Sound, Alaska. *Conservation Genetics*, 13, 1483–1497.

Toussaint R K, Sage G K, Talbot S L, Scheel D (2012b) Microsatellite marker isolation and development for the giant Pacific octopus (*Enteroctopus dofleini*). *Conservation Genetics Resources*, 4, 545-548.

Vaz-Pires P, Seixas P, Barbosa A (2004) Aquaculture potential of the common octopus (*Octopus vulgaris* Cuvier, 1797): a review. *Aquaculture*, 238, 221–238.

Vidal E A, Villanueva R, Andrade J P, Gleadall I G, Iglesias J, Koueta N, Rosas C, Segawa S, Grasse B, Franco-Santos R M, Albertin C B, Caamal-Monsreal C, Chimal M E, Edsinger-Gonzales E, Gallardo P, Le Pabic C, Pascual C, Roumbedakis K, Wood J (2014) Cephalopod culture: current status of main biological models and research priorities. *Advances in Marine Biology*, 67, 1–98.

Vieira M L C, Santini L, Diniz A L, Munhoz C D F (2016) Microsatellite markers: what they mean and why they are so useful. *Genetics and Molecular biology*, 39, 312–328.

Villanueva R, Sánchez P, Compagno Roeleveld M A (1991) *Octopus magnificus* (Cephalopoda: Octopodidae), a new species of large octopod from the southeastern Atlantic. *Bulletin of Marine Science,* 49, 39–56.

Villanueva R, Vidal E A G, Fernández-Álvarez F, Nabhitabhata J (2016) Early mode of life and hatchling size in cephalopod molluscs: Influence on the species distributional ranges. *PLoS One*, 11, e0165334.

Wakkanai Fisheries Research Institute (2022) Giant Pacific octopus in the waters around Hokkaido. In: Summary report on the stock status of giant Pacific octopus in the waters surrounding Hokkaido. *Hokkaido Research Organization Fisheries Research Department.* pp. 480–485. (in Japanese)

Waples R S (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, 89, 438–450.

Warnke K, Söller R, Blohm D, Saint-Paul U (2004) A new look at geographic and phylogenetic relationships within the species group surrounding *Octopus vulgaris* (Mollusca, Cephalopoda): Indications of very wide distribution from mitochondrial DNA sequences. *Journal of Zoological Systematics and Evolutionary Research*, 42, 306–312.

White C, Selkoe K A, Watson J, Siegel D A, Zacherl D C, Toonen R J (2010) Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences*, 277, 1685–1694.

Willing E M, Dreyer C, Van Oosterhout C (2012) Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. *PLoS One*, 7, e42649.

Wilson G A, Rannala B (2003) Bayesian inference of recent migration rates using multilocus genotypes. *Genetics,* 163, 1177–1191.

Winkelmann I, Campos P F, Strugnell J, Cherel Y, Smith P J, Kubodera T, Allcock, L, Kampmann M L, Schroeder H, Guerra A, Norman M, Finn J, Ingrao D, Clarke M, Gilbert M T P (2013) Mitochondrial genome diversity and population structure of the giant squid *Architeuthis*: genetics sheds new light on one of the most enigmatic marine species. *Proceedings of the Royal Society B: Biological Sciences*, 280, 20130273.

Wolf J B, Ellegren H (2017) Making sense of genomic islands of differentiation in light of speciation. *Nature Reviews Genetics*, 18, 87–100.

WoRMS Editorial Board (2025) World Register of Marine Species. Available from https://www.marinespecies.org at VLIZ. [Accessed September 4, 2025]

Wright S (1943) Isolation by distance. *Genetics*, 28, 114–138.

Wülker G (1910) Über Japanische cephalopoden: Beiträge zur kenntnis der systematik und anatomie der dibranchiaten. Abh Math-Phys Kl, K Bayer Akad Wiss. III. Suppl.--Bd. 1. *Abhandig*, 3, 1–77. (in German)

Yamada M, Ishibashi R, Toyoda K, Kawamura K, Komaru A (2014) Phylogeography of the brackish water clam *Corbicula japonica* around the Japanese Archipelago inferred from mitochondrial COII gene sequences. *Zoological Science*, 31, 168–179.

Yamazaki D, Miura O, Ikeda M, Kijima A, van Tu D, Sasaki T, Chiba S (2017) Genetic diversification of intertidal gastropoda in an archipelago: the effects of islands, oceanic currents, and ecology. *Marine Biology*, 164, 184

Appendix A. Numbers of samples with the genes and haplotypes at each site shown in Figures 3 and GenBank accession numbers for sequences used in this study for the cytochrome oxidase I gene (COI). Some haplotypes were integrated in other haplotype because of trimming. The total number of under column showed the number of specimens and (haplotypes).

COI

	Haplotype	Accession number	Soya	Obira	Mombetsu	Hidaka	Iburi	Mori	Shiriya	Asamushi	Fukui	Wakkanai	Iwate	Ishikawa	South Korea	TOTAL
NWP	EdCO_1	LC807055	31	25	43	13	20	12	20	1	2		1		2	170
	EdCO_2	LC807057	13	12	15	9	8	7	15						1	80
	EdCO_3	LC807056	1			1				1						3
	EdCO_4	LC807060	1			1		1								3
	EdCO_5	LC807068	1													1
	EdCO_6	LC807067	1													1
	EdCO_7	LC807066		1												1
	EdCO_8	LC807064		1	1											2
	EdCO_9	LC807061				1										1
	EdCO_10	LC807058				1										1
	EdCO_12	AB430536					1					1				2
	EdCO_15	AB191272												1		1
	TOTAL		48 (6)	39 (4)	59 (3)	26 (6)	29(3)	20 (3)	35 (2)	2 (2)	2 (1)	1 (1)	1 (1)	1 (1)	3 (2)	266 (12)

	Haplotype	Reference haplotype	Accession number	Dutch Harbor	Montague Strait	Kachemak Bay	Prince William Sound	Stephen's Passage	Glacier Bay	Canada	TOTAL
NEP	EdCO_9	CO01	KT2133445	1	1	4	2				8
		Haplotype1	JF306025	45		45	3	2	33		128
		EdCO_9	LC807061							1	1
	EdCO_12	EdCO_12	AB430536							1	1
	EdCO_16	Haplotype2*	JF306150				13	37			50
		CO02*	KT2133446				6				6
		CO03*	KT2133447				2				2
	EdCO_17	Haplotype3*	JF306187					2			2
		TOTAL		46 (1)	1 (1)	49 (1)	26 (1+1)	41 (1+2)	33 (1)	2 (2)	198 (2+2)

(*) Toussaint et al. (2012a) and Barry et al. (2013) suggested that the samples with these haplotypes were cryptic species.

Appendix B. Numbers of samples with the genes and haplotypes at each site shown in Figures 3 and GenBank accession numbers for sequences used in this study for the cytochrome *b* gene (CYTB). Some haplotypes were integrated in other haplotype because of trimming. The total number of under column showed the number of specimens and (haplotypes).

CYTB

	Haplotype	Accession number	Soya	Obira	Mombetsu	Hidaka	Iburi	Mori	Shiriya	Asamushi	Fukui	South Korea	TOTAL
NWP	EdCB_1	LC807721	47	34	57	24	28	19	34	2	2	3	250
	EdCB_2	LC807722	1		1								2
	EdCB_4	LC807724		2			1						3
	EdCB_5	LC807725		1	1								2
	EdCB_6	LC807726		1									1
	EdCB_7	LC807727			1								1
	EdCB_9	LC807729				1							1
	EdCB_10	KT213451				1							1
	EdCB_12	LC807733							1				1
	TOTAL		48 (2)	38 (4)	60 (4)	26 (3)	29 (2)	19 (1)	35 (1)	2 (1)	2 (1)	3 (1)	262 (9)

	Haplotype	Referencehaplotype	Accession number	Dutch Harbor	Montague Strait	Kachemak Bay	Prince William Sound	TOTAL
NEP	EdCB_10	CB04	KT213451	1		7	4	13
	EdCB_15	CB01	KT213448	8	4	21	11	44
	EdCB_16	CB02	KT213449			2	3	5
	EdCB_17	CB03	KT213450	8	1	23	14	46
	EdCB_18	CB05	KT213452			1		1
	EdCB_19	CB06	KT213453	1				1
	EdCB_20	CB07	KT213454	2				2
	EdCB_21	CB09	KT213456			1		1
	EdCB_22	CB10	KT213457		1	2		3
	EdCB_23	CB12	KT213459			1		1
	EdCB_24	CB08*	KT213455				1	1
	EdCB_25	CB11*	KT213458				8	8
		TOTAL		20 (5)	6 (3)	58 (8)	41 (4+2)	126 (10+2)

(*) Toussaint et al. (2012a) suggested that the samples with these haplotypes were cryptic species.

Appendix C1. Details of samples data and filtering for raw data to STACKS (Fukui, Hidaka, Iburi, Mombetsu, and Mori) (ddRAD-seq).

Collection site	Collection number	Sex	Haplotype CO1	Haplotype Cytb	Raw reads (bp)	fastp trimming q>33, ovaerrepresented > 1%, align to 100 bp	STACKS ref_map.pl ref. E. dofleini: GCA_027564315	STACKS population (-R 0.1 --max-orbs-het 0.6 --min-man 0.01 -write-random-snp	Frequency of missing (0 to 1)
Asamushi	21**	M	3	1	800,672	691,352	15,669	6,844	0.3024
Fukui	6**	paralarvae	1	1	1,634,484	1,406,566	19013	7,941	0.1906
	22* **	paralarvae	1	1	966,722	843,730	8,801	3,017	0.6925
Hidaka	233	M	2	1	955,842	834,777	16,244	6,956	0.291
	246	M	1	1	915,698	797,460	16,042	6,885	0.2982
	247	F	4	1	703,349	699,255	15,372	6,718	0.3153
	248	F	1	1	804,553	597,835	14,683	6,563	0.3311
Iburi	291	F	1	1	801,914	693,682	15,187	6,602	0.3271
	294	M	1	1	729,882	629,269	14,901	6,556	0.3318
	295	F	2	1	675,740	584,559	14,687	6,366	0.3511
	303	F	1	1	790,316	682,584	15,042	6,511	0.3364
Mombetsu	110	M	1	1	903,881	783,699	15,930	6,901	0.2966
	114	F	2	1	818,461	707,507	15,702	6,779	0.309
	117	F	1	1	832,604	714,275	15,679	6,720	0.3151
	119	F	8	1	747,764	643,666	15,306	6,639	0.3233
	145	F	1	1	1,039,742	901,321	16,233	7,008	0.2857
Mori	67	M	1	1	1,003,523	870,701	17,184	7,304	0.2555
	72	M	2	1	719,027	618,721	18,312	7,723	0.2128
	171	M	1	1	1,001,814	860,797	16,472	7,114	0.2749
	174	F	1	1	1,371,295	1,177,128	15,243	6,661	0.3211

Appendix C2. Details of samples data and filtering for raw data to STACKS (Obira, Shiriya, and Soya) (ddRAD-seq).

Collection site	Collection number	Sex	Haplotype CO1	Haplotype Cytb	Raw reads (bp)	fastp trimming trim: q>33, ovaerrepresented > 1%, align to 100 bp	STACKS ref_map.pl ref. E. dofleini: GCA_027564315	STACKS population (-R 0.1 --max-orbs-het 0.6 --min-man 0.01 -write-random-snp	Frequency of missing (0 to 1)
Obira	251	F	1	1	1,145,941	996,736	16,958	7,236	0.2625
	263	M	2	1	697,138	601,940	14,313	6,210	0.367
	274	F	8	1	781,135	676,405	15,670	6,765	0.3105
	280	F	1	1	848,141	734,778	15,405	6,565	0.3309
Shiriya	80	F	1	1	794,981	683,556	15,475	6,749	0.3121
	86	M	1	1	785,631	679,511	15,141	6,624	0.3248
	88	F	2	1	658,419	568,661	14,916	6,497	0.3378
	92	M	2	1	713,660	614,515	15,199	6,598	0.3275
Soya	185	M	1	1	777,975	676,221	15,739	6,687	0.3184
	188	F	2	1	873,168	758,350	15,648	6,653	0.3219
	198	M	1	1	694,328	600,905	15,158	6,628	0.3244
	203* **	M	5	2	1,097,691	955,666	13,812	5,367	0.453
Sum (all sites)					9,868,208	8,547,244	183434 sites	78579 sites	
average					822,351	712,270	Total 266434 loci	Total 60278 loci	
variant sites							30,856	9,811	
genotyping rate		(analysed loci/total loc)						0.688398	

Appendix C3. Details of samples data and filtering of two datasets (Fukui, Hidaka, Iburi, Mombetsu, and Mori) (ddRAD-seq).

		Dataset IND (n=30)			Dataset POP (n=28)		
Collection site	Collection number	PLINK (--maf 0.05 --geno 0.2	PCAdapt K=16 FDA 0.01	Frequency of missing (0 to 1)	PLINK (--maf 0.05 --geno 0.4	PCAdapt K=15 FDA 0.01	Frequency of missing (0 to 1)
Asamushi	21**	1626	1435	0.05963	-	-	-
Fukui	6**	1682	1482	0.02883	-	-	-
	22* **	-	-	-	-	-	-
Hidaka	233	1638	1445	0.05308	2177	1847	0.1316
	246	1625	1433	0.06094	2133	1806	0.1509
	247	1609	1426	0.06553	2115	1795	0.1561
	248	1582	1400	0.08257	2049	1736	0.1838
Iburi	291	1596	1419	0.07012	2086	1772	0.1669
	294	1587	1402	0.08126	2043	1740	0.1819
	295	1575	1392	0.08781	1987	1684	0.2083
	303	1560	1377	0.09764	2021	1703	0.1993
Mombetsu	110	1632	1442	0.05505	2177	1850	0.1302
	114	1617	1430	0.06291	2097	1786	0.1603
	117	1611	1419	0.07012	2092	1776	0.165
	119	1595	1409	0.07667	2060	1761	0.1721
	145	1639	1441	0.0557	2183	1857	0.1269
Mori	67	1655	1464	0.04063	2186	1865	0.1232
	72	1588	1401	0.08191	2046	1735	0.1843
	171	1654	1461	0.0426	2225	1889	0.1119
	174	1688	1489	0.02425	2285	1940	0.08792

* were excluded from Dataset IND. ** were excluded from Dataset POP.

Appendix C4. Details of samples data and filtering for two datasets (Obira, Shiriya, and Soya) (ddRAD-seq).

		Dataset IND (N=30)			Dataset POP (n=28)		
		PLINK	PCAdapt		PLINK	PCAdapt	
Collection site	Collection number	(--maf 0.05 --geno 0.2	K=16 FDA 0.01	Frequency of missing (0 to 1)	(--maf 0.05 --geno 0.4	K=15 FDA 0.01	Frequency of missing (0 to 1)
Obira	251	1651	1458	0.04456	2210	1883	0.1147
	263	1548	1376	0.0983	1972	1675	0.2125
	274	1618	1427	0.06488	2107	1788	0.1594
	280	1591	1404	0.07995	2044	1735	0.1843
Shiriya	80	1595	1407	0.07798	2079	1763	0.1711
	86	1598	1414	0.07339	2068	1763	0.1711
	88	1590	1404	0.07995	2038	1729	0.1871
	92	1578	1392	0.08781	2037	1729	0.1871
Soya	185	1575	1396	0.08519	2073	1770	0.1678
	188	1591	1404	0.07995	2063	1742	0.181
	198	1564	1384	0.09305	2043	1729	0.1871
	203* **	-	-	-	-	-	-
Sum (all sites)							
average			202 outliers			357 outliers	
variant sites		1728	1526		2484	2127	
genotyping rate		0.930903	0.931258		0.843915	0.836994	

* were excluded from Dataset IND. ** were excluded from Dataset POP.

Appendix D. Details of the seven microsatellite markers used in this study based on Toussaint et al. (2012b). The forward primers were used with each tailed universal primer and fluorescent array.

Locus name	Repeat motif		Primer sequence (bp)	Size range	GenBank accession number
Edol4	$(CA)_{12}$	F	M13-AGCCCTCGAGCATATTCG	148–168	JQ236863
		R	gtttcttTGATGTGGTTCCAGTGTC		
Edol6	$(CA)_4G(CA)_{23}(TA)_2CATA(CA)_9$	F	PQE-E-TTAGTTGGATTGCTGGGC	180–220	JQ236864
	$(TG)(CA)_5(TA)(CA)_2$	R	gtttcttCAGTGGATCAACAATAGC		
Edol8	$(CA)_2(GA)(CA)_2CGACAAATA$	F	pVP16-GCCAATCACAAATAGTGAC	179–233	JQ236865
	$(CA)_{27}AG(CA)_5AG(CA)_3$	R	gtttcttCTACTTAGGTGTACGTGC		
Edol115	$(CA)_{18}$	F	RV3-CAGGGAAACCGAACTTGC	172–186	JQ236871
		R	gtttcttATACGTACAACTGCTGGC		
Edol118	$(CA)_{28}$	F	M13-GGCGACAGTGATCATAGC	225–269	JQ236873
		R	gtttcttTTTTCATACCCATCCAGC		
Edol123	$(CA)_{15}$	F	M13-GATATGCAACATCTCAGC	167–177	JQ236880
		R	gtttcttCAGTACGGAGTTAGTGGC		
Edol216	$(CA)_{10}(TA)_4$	F	pVP16-TACGATTGTAAATCCGGTAGC	113–125	JQ236878
		R	gtttcttGTATAATCACCATCATCG		

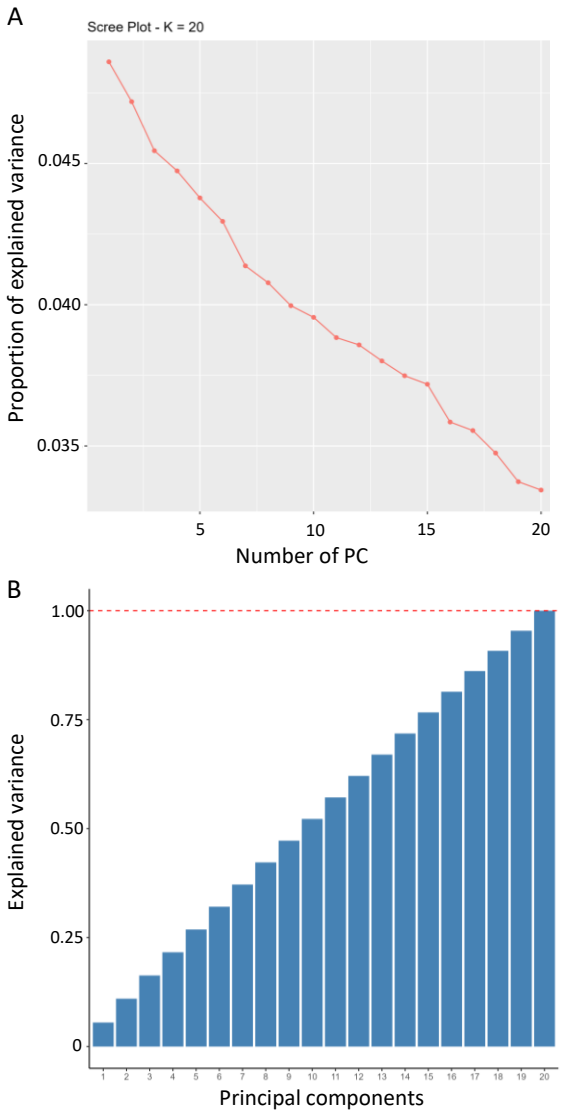
(6-PET) M13: TGTAAAACGACGGCCAGT

(NED) PQE-E: TTGAGAGGATCGCATCCA

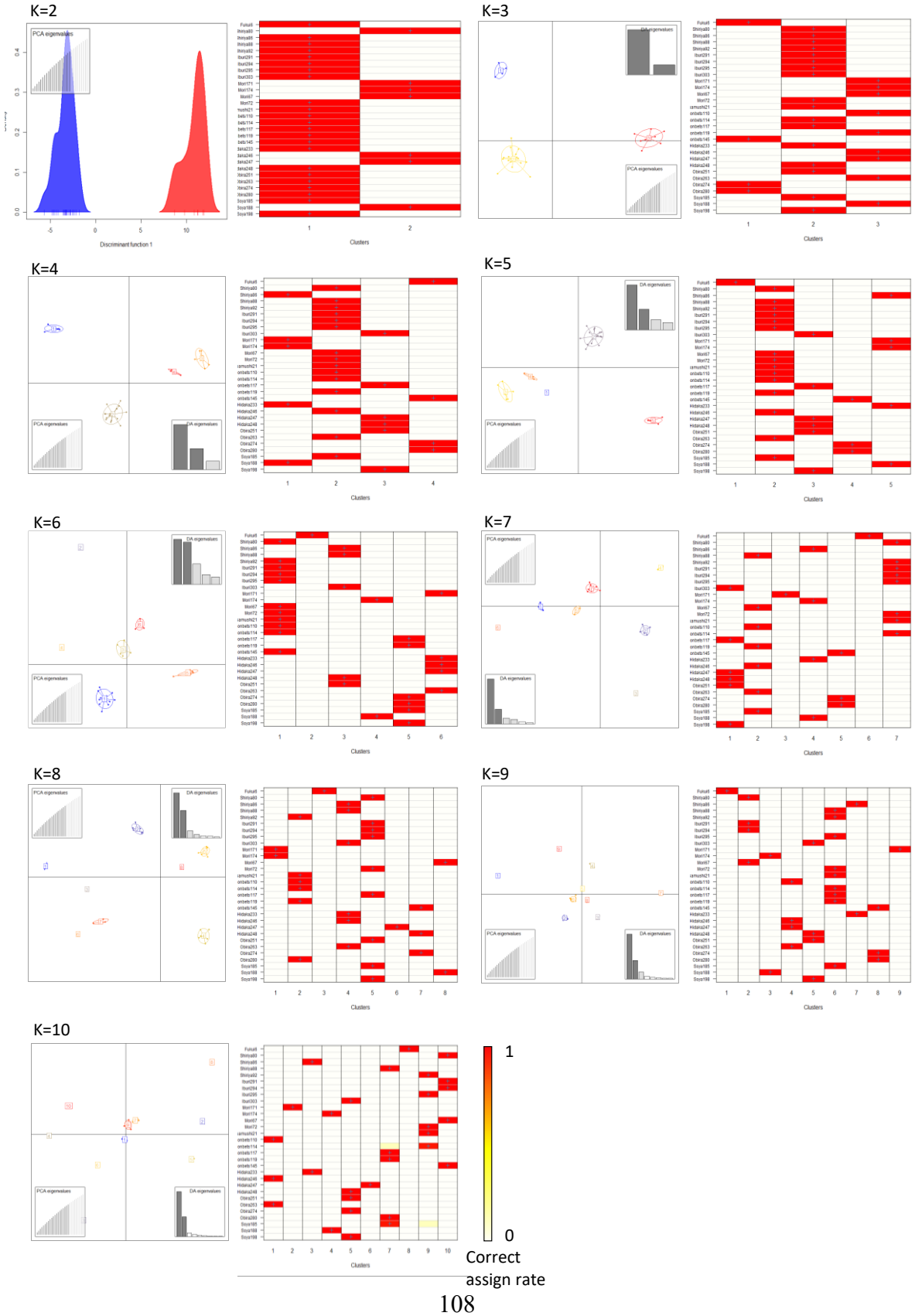
(VIC) pVP16: GCCGACTTCGAGTTTGAG

(FAM) RV3: AGCAAAATAGGCTGTCCC

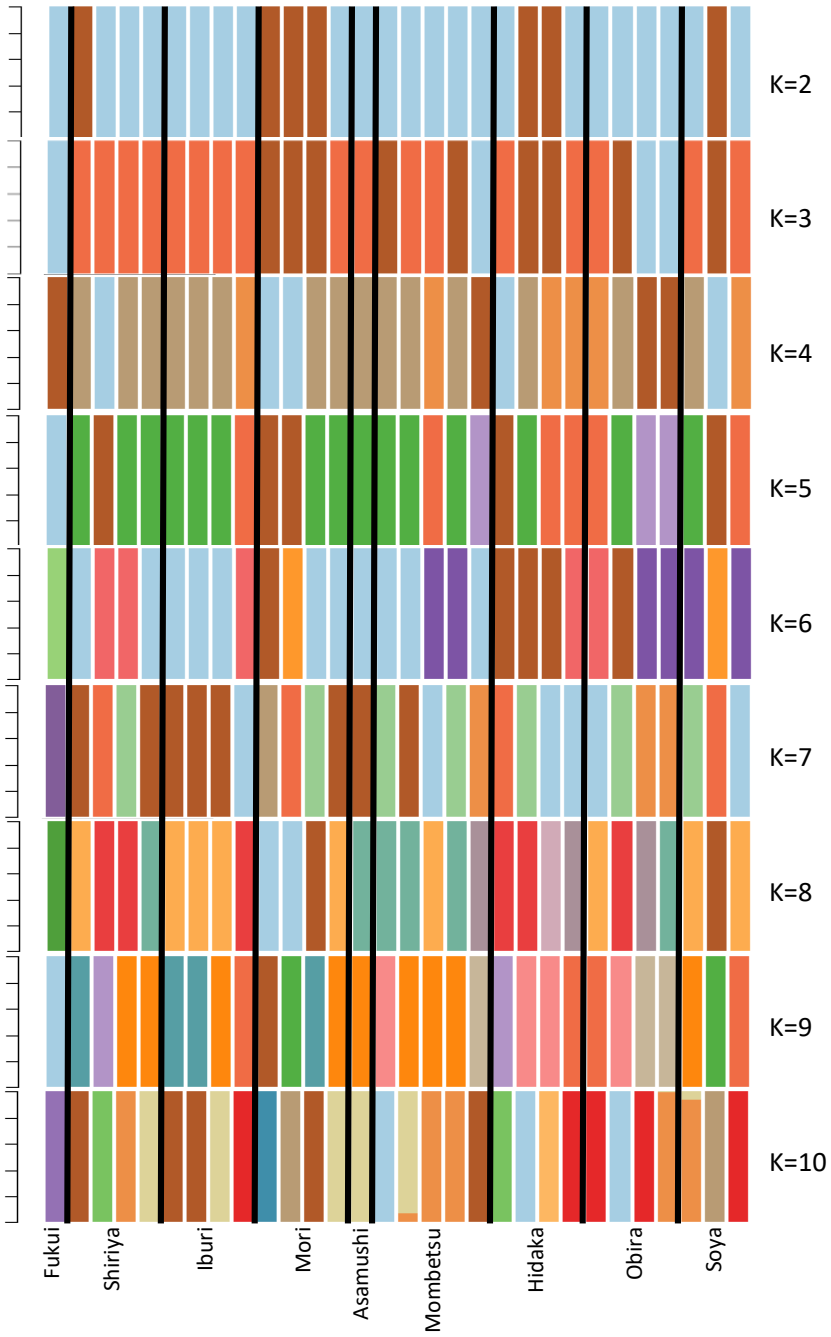
Appendix E. Scatterplot and explained variance for the PCA with Dataset POP. A; Scree plot to determine the appropriate PC numbers to retain. B; Explained variance for each PC.



Appendix F. Assign plot and scatter plot of DAPC (K = 2 to 10). Left panel is the scatter plot of DAPC and right panel is the assignment plot. Assignment accuracy is indicated by color, with red representing high assignment probability and white representing low.



Appendix G. Component plot of DAPC (K = 2 to 10). The sampling sites are shown at the bottom, with samples colored according to the cluster they belong to. However, the colors for different values of K are not consistent across clusters.



Appendix H. Accession no. and length (bp) of the downloaded sequences to construct pylogenetic tree.

Species	Region	Accession num.	Length (bp)
Argonauta nodosa	complete	NC048475	15201
Bathypolypus sponsalis	complete	NC069190	17207
Enteroctopus dofleini	complete	NC056385	16026
Muusoctopus johnsonianus	complete	NC069994	16306
Muusoctopus leioderma	complete	NC069994	16408
Octopus bimaculatus	complete	NC028547	16084
Octopus vulgaris	complete	NC006353	15744
Opisthoteuthis grimaldii	complete	NC069991	16465
Todarodes pacificus	complete	NC006354	20254
Tremoctopus violaceus	complete	NC070289	16191
Vampyroteuthis infernalis	complete	NC009689	15617
Enteroctopus zealandicus	COI	MT225042	500
	COIII	MT216543	638
Enteroctopus megalocyathus	COI	KM459468	657
	COIII	KM459484	527
	Cytb	HM572210	434
Vulcanoctopus hydrothermalis	COI	HM104264	657
	COIII	FJ603533	529

Appendix I. Phylogenetic tree of 14 cephalopods based on Bayesian inference of 13 protein coding regions from the mitochondrial gene with *Todarodes pacificus* as the outgroup. Divergence time estimates on X-axis are shown in millions of years. Blue bars show 95% CI. The Bayesian posterior probabilities and Bootstrap scores of the clades were shown below the nodes (posterior probability/ Bootstrap score). Only bootstrap values > 65 and posterior probabilities > 0.9 are shown. The bootstrap values = 100 and posterior probabilities = 1 are shown as *.

