**Population genomics of blue shark *Prionace glauca* in the Pacific Ocean based on whole mitogenome sequencing and nuclear-genome-wide SNP data**

Masayoshi Ibe\*1, Shohei Kokubun\*1, Izumi Funatogawa\*1, Keigo Tsumita\*1, Shigehiro Kuraku\*2, Jacquelynne King\*3, John R. Hyde\*4, Yuki Fujinami\*5, Akira Kurashima\*6, Mikihiko Kai\*6, Kenji Nohara\*1

\*1 TOKAI University, Department of Marine Biology, School of Marine Science and

Technology. 3-20-1, Orido, Shimizu, Shizuoka 424-8610, Japan

\*2 Laboratory for Phyloinformatics, RIKEN Center for Biosystems Dynamics Research,

2-2-3 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo, 650-0047, Japan

Present address:

Molecular Life History Laboratory, National Institute of Genetics, Mishima, Shizuoka, 411-8540, Japan

\*3 Pacific Biological Station, Fisheries and Oceans Canada, Nanaimo, British Columbia

V9T 6N7

\*4 Fisheries Resources Division, Southwest Fisheries Science Center, National Marine

Fisheries Service, National Oceanic and Atmospheric Administration, La Jolla, CA 92037,

United States

\*5 Fisheries Stock Assessment Center, Fisheries Resources Institute, Japan Fisheries and

Education Agency, 1551-8 Taira-machi, Nagasaki, Nagasaki 851-2213, Japan

\*6 Fisheries Stock Assessment Center, Fisheries Resources Institute (FRI), Japan Fisheries and Education Agency, 2-14-4 Fukuura, Kanazawa, Yokohama, Kanagawa, 236-8648, Japan

**Abstract**

Blue shark *Prionace glauca* is highly migratory pelagic fish, inhabit all of the world's oceans, except for the polar area. The information of tagging data and reproductive ecology do not support panmixia of blue shark stock in the Pacific Ocean, but genetic data in previous studies could not reject the null hypothesis of a single panmictic population in the Pacific Ocean. This conflict may be resolved by contemporary genome-wide approach such as genotyping-by sequencing approach (GBS). In this paper, we report about population genomics for blue shark in the Pacific Ocean using whole mitochondrial genome sequencing, Genotyping by Random Amplicon Sequencing-Direct (Gras-Di) for nuclear genome, and *de novo* whole genome assembly. We evaluated genetic variation from whole mitochondrial genome sequencing data (16,705 base-pair long) and 3,816 genome wide SNPs for a total of 280 specimens collected from eight sampling sites. Two major haplotype clades (Clade1 and Clade2) and sub-clade were revealed from whole mitochondrial genome sequence data, but heterogeneity of haplotype frequency among sampling sites was not found in the Pacific Ocean. On the other hand, our genome-wide SNP data showed that some east Pacific individuals were separated from the other major group. Therefore, possibility of genetic differentiation may exist between east and west Pacific rather than north-south Pacific. To obtain the more informative SNPs, we are working on low-coverage whole genome resequencing (lcWGS) and genome assembly by chromosome-scale scaffolding.

**1. Introduction**

Blue shark *Prionace glauca* is a highly migratory pelagic fish species that inhabits all of the world's oceans, except for the polar area (Compagno, 1984; Nakano & Seki, 2003; Queiroz et al., 2012). Stock assessment of this species has traditionally been conducted by four Regional Fisheries Management Organizations (RFMOs); in the Atlantic Ocean (ICCAT: The International Commission for the Conservation of Atlantic Tunas), in the Indian Ocean (IOTC: The Indian Ocean Tuna Commission), in the North Pacific Ocean (ISC: The International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean) and South Pacific Ocean (WCPFC: The Western and Central Pacific Fisheries Commission), and these stocks were managed by ICCAT, IOTC and WCPFC. For the sustainable use of this species, it is important to clarify the stock structure considering not only stock boundary delineated by these geopolitical frameworks but also biological factors such as migratory patterns, reproductive ecology and genetic information.

Several tagging studies for blue sharks indicated clear spatial segregation by sex and life-history stages (Howey et al., 2017; Maxwell et al., 2019; Fujinami et al., 2022). In the Pacific Ocean, although broad-scale movements over thousands of kilometers and complex migration patterns have been reported for blue sharks based on conventional and electronic tagging studies (Musyl et al., 2011; Maxwell et al., 2019), those studies didn’t clearly show east-west active movements such as Pacific bluefin tuna (Fujioka et al., 2018) and trans-equatorial migration (Sippel et al., 2011; Musyl et al., 2011; Maxwell et al., 2019; Fujinami et al., 2022,). In addition, non-overlapping reproductive cycles have been reported in the Northern and Southern Hemispheres (Nakano & Seki, 2003; Nakano & Stevens, 2008). Thus, the information of tagging data and reproductive ecology do not support panmixia of blue shark stock in the Pacific Ocean.

Previous genetic studies of blue sharks in the Pacific Ocean, which are conducted based on partial mitochondrial sequence data (Taguchi et al., 2015) and 14 microsatellite DNA markers (King et al., 2015), could not reject the null hypothesis of a single panmictic population in the Pacific Ocean. Most highly migratory marine fish species typically have large effective population sizes (Ne), which is panmictic populations with frequent dispersal and high gene flow (Nielsen et al., 2009; Waples et al.,2008). In fact, the blue shark was used as a case species to illustrate the concept of “population grey zone” (Bailleul et al., 2018). However, recent genome-wide single nucleotide polymorphism (SNP)-based approach using DrTseq (Diversity Arrays Technology sequencing; Jaccoud et al., 2001) clearly identified different population genetic structure between north Atlantic and south Atlantic-Indian Ocean (Nikolic et al., 2022). Although a large number of evidences are accumulated to demonstrate that the genome-wide approach is effective in revealing fine-scale population genetic structure for highly migratory fishes (e.g., Albacore; Vaux et al., 2021), the population genomics for blue sharks in the Pacific Ocean has not been adopted and fine-scale population structure remained unclear. In this paper, we report population genomics of blue shark in the Pacific Ocean using whole mitogenome sequencing, genotyping-by sequencing (GBS) approach called Gras-Di; Genotyping by Random Amplicon Sequencing-Direct, and *de novo* whole genome assembly.

**2. Material and Methods**

2-1. Sample collection and DNA extraction

288 individuals were collected by research vessels from eight Pacific sampling sites between 1995 and 2019 (Table.1, Fig.1): off Choshi, Japan (CHS), off Sanriku, Japan (SRK), off Hawaii, USA (HI), off southern California, USA (SCA), off Peru (PER), off Tasmania, Australia (TAS), off Chile (CHI), off British Columbia, Canada (BC). DNA extraction from muscle or fin tissues was conducted by DNeasy Bood & Tissue Kit (QIAGEN) with RNase A (QIAGEN), and the assessment 　of quality and quantity of the genomic DNA were checked by 1.0 % agarose gel electrophoresis and Qubit 4 Fluorometer (Thermo Fisher Scientific), respectively.

2-2. Draft genome assembly

High-molecular-weight DNA was extracted from muscle of the individual #138 G-7 using a NucleoBond AXG column (Macherey-Nagel), which was followed by purification with phenol-chloroform. The concentration of the extracted DNA was measured with Qubit (ThermoFisher), and their size distribution was first analyzed with TapeStation 4200 (Agilent Technologies) to ensure high integrity and later analyzed with pulse-field gel electrophoresis on CHEF DR-II (BioRad) to ensure the size range between 20 kb to 100 kb. To obtain long-read sequence data, a SMRT sequence library was constructed with an SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, Menlo Park, CA, USA) and was sequenced in five 8M SMRT cells on a PacBio Sequel II system (Pacific Biosciences). The sequencing output was processed to generate circular consensus sequences (CCS) to obtain a total of 148 Gb HiFi sequence reads. From these reads, adapter sequences were removed using the program HiFiAdapterFilt (Sim et al. 2022). The obtained HiFi sequence reads were assembled using the program hifiasm v0.15.5 (Cheng et al. 2021) with its default parameters, and the output was designated as ‘v1’ assembly. The N50 scaffold length of this assembly amounted to 8.57 Mbp, and the BUSCO ‘Complete’ score of 1-to-1 ortholog coverage reached 94%, which ascertains the high continuity and completeness of the assembly. Repetitive sequences in the assembly v1 were detected and hard-masked by RepeatMasker v4.0.1 (http://www.repeatmasker.org) with a species-specific repeat library constructed by RepeatModeler v2.0.1 (https://www.repeatmasker.org/RepeatModeler/), which is followed by more thorough masking of simple tandem repeats detected by TRF v4.09.1 (Benson 1999). The output hard-masked assembly is designated as ‘v1.hm’ assembly.

2-3. NGS library prep. and data processing for population genomics

2-3-1. Mitogenome

Genomic DNA libraries were prepared using Illumina DNA Prep (Illumina, Inc.) according to the HackFlex protocol (Gaio et al., 2022). Then, NGS Sequencing was performed on an Illumina HiSeq X Ten (150 bp paired-end) with a target acquisition data volume of 1.0 Gb/individuals for constructing whole mitochondrial genome data.

Sequencing reads were quality checked using the FastQC version 0.11.9 quality control tool (Andrews, 2010), and low-quality reads were trimmed using Fastp version 0.22.0 (Chen et al., 2018). Trimmed sequence reads were mapped on the whole mitochondrial sequence data of the blue shark (NC\_022819, Chen et al., 2013) using BWA-MEM2 version 2.2.1(BWA-MEM2. https://github.com/bwa-mem2/bwa-mem2). Mapped sequence reads were sorted by Samtools version 1.11 (Li et al., 2009), and duplicate reads were removed using picard tools version 2.27.4 (Broad Institute, 2019; Picard, https:// broadinstitute.github.io/picard/). Variants were called using GATK version 4.2.6.1 (Genome Analysis Toolkit; DePristo et al., 2011) HaplotypeCaller. Variant filtrations were performed GATK hardfilter and Vcflib 1.0.3 (Garrison et al., 2022) filtering. Finally, consensus sequence of whole mitogenome was generated using R script (R Core Team, 2018). In addition, *de novo* assembly was also conducted by GetOrganelle version 1.7.1 (Jin et al., 2020) to check the consensus sequence based on the mapping pipeline. All consensus sequences were aligned by using MAFFT version 7.52 (Katoh & Standley, 2013), and mitogenome genes were annotated by MitoAnnotator (<http://mitofish.aori.u-tokyo.ac.jp/annotation/input/>) and visualized by ShinyCircos software (Yu et al., 2018).

2-3-2. GRAS-Di analysis

281 samples of blue sharks from 8 localities were used for genotyping by random amplicon sequencing-direct (GRAS-Di) (Hosoya et al., 2019). Library preparation and using the GRAS-Di method were outsourced to GeneBay Co., Ltd (Yokohama, Japan) and the library preparation protocols were followed by Hosoya et al. (2019). The random amplicons were produced with 64 primers and sequenced on a MGISEQ-G2000RS sequencer (MGI Tech Co., Ltd), with a 150 bp-paired end (PE) read length. Low-quality bases and Illumina sequencing adapters were trimmed using the Fastp version 0.22.0. Sequences of average quality <20 and reads shorter than 50 bp were removed. Trimmed Gras-Di reads were mapped on to the blue shark genome assembly v1.hm (see above) using BWA-MEM2 version 2.2.1.

　 For SNP calling and genotyping, we used the `ref\_map.pl` pipeline with v1.hm assembly and `Populations` pipeline in Stacks version 1.47 (Catchen et al., 2011, 2013). *Pstacks, cstacks, sstacks* and *populations* programs are executed in the `ref\_map.pl` and `Populations` pipeline. The *pstacks* (parameter; -m 10) was used to assemble loci according to the alignment positions provided for each read and calling SNPs in each sample. The *cstacks* was used to create catalog of all loci across the population. Loci from different samples are matched up across the data set according to alignment position. The *sstacks* was used to match each sample against the catalog created in *cstacks*. The *populations* was used to generate population-level summary statistics with specified population map (-O option) which set with each sampling point. We also used the *populations* program to identify loci between sampling sites which required 20 individuals at a locus (-m 20); only the first SNP in each locus was used for downstream analyses to maintain independence of loci (--write-single-snp). Finally, the dataset was filtered by using PLINK version 1.90b6.21 (Chang et al., 2015), with the following parameters: a genotype call rate per SNP higher than 90% (-geno 0.1), a missing genotype call rate per sample lower than 40% (-mind 0.4), and a threshold for minor allele frequency (MAF) of 1% (-maf 0.01).

2-4. Data analyses

2-4-1. Mitogenome

Genetic diversity indices (nucleotide diversity; π and haplotype diversity; *h*) and genetic differentiation (F-statistics; *F*ST) were calculated and then Analysis of Molecular Variance (AMOVA) was performed using Arlequin version 3.5.2.2 (Excoffier & Lischer, 2010). Neighbor-joining tree (K2P model) and minimum spanning network of blue shark haplotypes in the Pacific Ocean were reconstructed by MEGA version 10.0.5 (Kumar et al., 2018) and POPART 1.7 (Leigh & Bryant, 2015), respectively. Mismatch distribution analysis based on pairwise differences among individuals, Tajima's*D* (Tajima 1989) and Fu's *Fs* (Fu 1997) were calculated using Arlequin.

2-4-2. Nuclear genome

We carried out population genetic analyses based on the filtered SNPs data set (3,816 SNPs) using principal component analysis (PCA), structure analysis by using ADMIXTUER version 1.3 software (Alexander et al., 2009) and principal coordinate analysis (PCoA) was conducted by GenAlEx version 6.5 (Peakall & Smouse 2006, 2012). The PCA command in PLINK version 1.90b6.21 was used to calculate the principal components for individual from the filtered SNPs data and to convert vcf file to bed file for structure analysis. R studio (https://posit.co/download/rstudio-desktop/) was then used to calculate the contribution of each principal component and to visualize the results. ADMIXTURE version 1.3. was run to assign populations to ancestral clusters and determine population relatedness. We ran ADMIXTURE from K = 1 to K = 6 and the most likely value of K was assessed by compared cross-validation error’s (cv error) values of each K. The principal coordinates analysis by GenAlEx was carried out based on the *F*ST values between each sampling site calculated by STACKS. The analysis process followed its official website (<https://biology-assets.anu.edu.au/GenAlEx/Welcome.html>). We detected outlier loci in the SNPs data set to confirm adaptive variation within the Pacific Ocean. Three software packages were used to detected outlier loci; BayeScan version 2.1 (Foll et al., 2012) for Bayesian approach, Outflank version 0.2 (Whitlock et al., 2015) for *F*ST-based outliers and pcadapt version 4.3.3(Luu et al., 2017) for outliers which individual-based structure. In BayeScan, prior odds set to 100 and other parameters set to default. Outflank was run with default options (LeftTrimFraction= 0.05, RightTrimFraction = 0.05, Hmin = 0.1) and the Q-threshold set to 0.01. In pcadapt, the number of Principal Components was set to 2 (K=2) and use three different cut-off method for outlier detection; q-values, Benjamini-Hochberg Procedure (Benjamini et al,.1995) and Bonferroni correction (Bonferroni., 1936) (p<0.05).

**3. Result and discussion**

For mitogenome analysis, we obtained 264 complete consensus mitogenome sequences (16,705 bp) from original mapping pipeline (68 x mean depth of coverage) and checked with *de novo* assembly dataset (Fig. 2). Overall, 540 variable sites were identified in 236 haplotypes. ND1 gene regions showed higher numbers of polymorphisms (π=0.4%) than previous study; cytb gene region (approx.0.2%, Taguchi et al., 2015; Leone et al., 2017) and control region (approx.0.3%, Leone et al., 2017; Veríssimo et al.,2017). Haplotype (h) and nucleotide diversities (π) of whole mitogenome sequences were similar among all sampling sites, ranging from 0.996 to 1.000 and from 0.13 to 0.18 % respectively (Table 1). Two major haplotype clades (Clade1 and Clade2) and sub-clade were revealed by phylogenetic tree (Fig.3 (A)) and haplotype network (Fig.3 (B)) analyses. These results were supported by mismatch distribution analysis based on the whole mitogenome sequences (Fig. 4). Pairwise *F*ST values did not show significant genetic differentiation among sampling sites after sequential Bonferroni correction (Table 2). Any genetic relationship was not found in mitogenome analyses among sampling sites within the Pacific Ocean.

For nuclear genome analysis, we obtained 72,195 Gras-Di-SNPs from 281 individuals before filtering. Ultimately, 3,816 Gras-Di-SNPs with hard-masked reference genome v1.hm were obtained from a total number of 280 individuals (one individual from CHS15 was eliminated after filtering). In the PCA analysis, the first, second and third PC accounts for 6.7 %, 6.4 % and 5.2% of the variation, respectively (Fig. 5(A), (B)). Ten individuals in east Pacific Ocean (SCA, BC, PER) and two individuals from northwestern Pacific Ocean (CHS) were distantly placed from the other major central plots. The ADMIXTURE analysis exhibited the lowest and second-lowest CV error at K = 1 and K = 2, respectively (Fig.6 (A)). The probability of each individual assigned to two genetic clusters (K=2) showed in Fig.6 (B). No apparent ancestral population cluster within the Pacific Ocean was detected. In the PCoA analysis based on the genetic distance (pairwise-*F*ST values) among sampling sites, there were no spatial-temporal genetic differences between north and south sampling sites such as CHS and TAS (Fig.7). On the other hand, north-east (SCA and BC) and south-east (CHL) populations were plotted in different positions than the others. Based on these results and no trans-equatorial migration reported in previous studies, we divide into four geographical areas to detect outlier loci using Gras-Di dataset; North Pacific (CHS, SRK, HI, BC, SCA), South Pacific (TAS, PER, CHL), Western Pacific (CHS, TAS), and Eastern Pacific (BC, SCA, PER, CHL). In both BayeScan and pcadapt analyses, four overlapped outlier loci (three loci on the scaffold 89 and one locus on the scaffold 944) were detected only between western and eastern Pacific Ocean.

Our results indicated that past genetic isolation and subsequent genetic mixing (secondary contact) inferred from whole mitochondrial genome data, and the possibility that genetic differentiation of nuclear genome may exist between east and west Pacific rather than north-south Pacific. Global-scale sampling and whole mitochondrial genome analysis would be essential to describe the evolutionary scenario for the origin of two major mitochondrial clades and secondary contact. The possibility of genetic differentiation between east-west Pacific may be explained by the following studies. First, most tracking data from conventional and electric tags show predominantly latitudinal movements (Musyl et al.,2011; Maxwell et al., 2019), namely trans-oceanic movement between north-west and north-east Pacific Ocean is limited. Second, stable isotope analysis (SIA)-based movement studies have also been supported limited east-west movement (Madigan et al.,2021). Although these results may support our nuclear genome analysis in the present study, a potential north-south and/or east-west genetic cline in the Pacific Ocean remains uncertain at present. We are therefore working on low-coverage whole genome sequencing (lcWGS) approach while the genome assembly is being improved by chromosome-scale scaffolding using chromatin contact data obtained by the Hi-C method to gain more informative SNPs than the reduced-representation sequencing approach (Gras-Di).

**4. Acknowledgements**

We thank all of the fishermen, scientific observers, and researchers who contributed to the collection of the samples. We also thank Dr. H. Takeshima and Dr. Y. Semba for valuable comments on the manuscript. This work was supported by a grant from the Fisheries Agency of Japan.

**5. References**

Alexander, DH., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 19, 1655-1664.

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

Bailleul, D., Mackenzie, A., Sacchi, O., Poisson, F., Bierne, N., and Arnaud-Haond, S. (2018). Large-scale genetic panmixia in the blue shark (Prionace glauca): A single worldwide population, or a genetic lag-time effect of the "grey zone" of differentiation? Evolutionary Applications, 11, 614-630.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological), 57, 289-300.

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Research, 27(2), 573-580.

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze, 8, 3-62.

Broad Institute. (2019). Picard toolkit. Broad Institute, GitHub repository. http://broadinstitute.github.io/picard/.

BWA-MEM2. https://github.com/bwa-mem2/bwa-mem2

Catchen, J., Amores, A., Hohenlohe, P., Cresko, W., and Postlethwait, J. (2011). Stacks: building and genotyping loci de novo from short-read sequences. G3: Genes, Genomes, Genetics, 1, 171-182.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., and 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., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience, 4, 7.

Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nature Methods, 18, 170–175.

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor. Bioinformatics, 34, i884-i890.

Chen, X., Xiang, D., Ai, W., and Shi, X. (2013). Complete mitochondrial genome of the blue shark Prionace glauca (Elasmobranchii: Carcharhiniformes). Mitochondrial DNA, 26(2), 313-314.

Compagno, L. J. V. (1984). FAO species catalogue, Vol. 4, Sharks of the World. An annotated and illustrated catalogue of shark species known to date (Part 2, Carcharhiniformes). FAO Fisheries Synopsis, 125(4), 655.

DePristo, M., Banks, E., Poplin, R., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics, 43, 491-498.

Excoffier, L., and Lischer, H. E. (2010). Arlequin Suite Ver 3.5: A New Series of Programs to Perform Population Genetics Analyses under Linux and Windows. MOLECULAR ECOLOGY RESOURCES, 10(3), 564-567. doi: 10.1111/j.1755-0998.2010.02847.x

Foll, M. (2012). BayeScan v2. 1 user manual. Ecology, 20, 1450-1462.

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

Fujinami, Y., Kurashima, A., Shiozaki, K., et al. (2022). New insights into spatial segregation by sex and life-history stage in blue sharks Prionace glauca in the northwestern Pacific. Mar. Ecol, 696, 69-84. https://doi.org/10.3354/meps14130.

Fujioka, K., Fukuda, H., Tei, Y., et al. (2018). Spatial and temporal variability in the trans-Pacific migration of Pacific bluefin tuna (Thunnus orientalis) revealed by archival tags. Progr. Oceanogr. 162, 52–65.

Gaio, D., Anantanawat, K., To, J., Liu, M., Monahan, L., and Darling, A. E. (2022). Hackflex: Low-Cost, High-Throughput, Illumina Nextera Flex Library Construction. Microbial Genomics, 8(1). doi: 10.1099/mgen.0.000744

Garrison, E., Kronenberg, Z. N., and Dawson, E. T. (2022). A Spectrum of Free Software Tools for Processing the VCF Variant Call Format: Vcflib, Bio-Vcf, Cyvcf2, Hts-Nim and Slivar. PLOS COMPUTATIONAL. doi: 10.1371/journal.pcbi.1009123

GenALEx. https://biology-assets.anu.edu.au/GenAlEx/Welcome.html

Hosoya, S., Hirase, S., Kikuchi, K., et al. (2019). Random PCR-based genotyping by sequencing technology GRAS-Di (genotyping by random amplicon sequencing, direct) reveals genetic structure of mangrove fishes. Mol. Ecol, 19, 1153–1163.

Howey, L. A., Wetherbee, B. M., Tolentino, E. R., and Shivji, M. S. (2017). Biogeophysical and physiological processes drive movement patterns in a marine predator. Mov. Ecol, 5, 16, https://doi.org/10.1186/s40462-017-0107-z.

Jaccoud, D., Peng, K., Feinstein, D., and Kilian, A. (2001). Diversity Arrays: A solid state technology for sequence information independent genotyping. Nucleic Acids Research, 29, e25.

Jin, J., Yu, W., Yang, J., et al. (2020). GetOrganelle: A Fast and Versatile Toolkit for Accurate de Novo Assembly of Organelle Genomes. Genome Biology, 21, 241. doi: 10.1186/s13059-020-02154-5

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol, 30, 772–780.

King, JR., Wetklo, M., Supernaul, J., et al. (2015). Genetic analysis of stock structure of blue shark (Prionace glauca) in the north Pacific. Fish Res, 172, 181–189

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Molecular Biology and Evolution, 35(6), 1547–1549. doi: 10.1093/molbev/msy096

Leigh, J. W., and Bryant, D. (2015). Popart: Full-Feature Software for Haplotype Network Construction. Methods in Ecology and Evolution, 6(9), 1110–1116. doi: 10.1111/2041-210X.12410

Leone, A., Urso, I., Damalas, D., et al. (2017). Genetic Differentiation and Phylogeography of Mediterranean-North Eastern Atlantic Blue Shark (Prionace Glauca, L. 1758) Using Mitochondrial DNA: Panmixia or Complex Stock Structure? Peerj, 5, 18.

Li, H., Handsaker, B., Wysoker, A., et al. (2009). The Sequence Alignment/Map Format and SAMtools. Bioinformatics, 25(16), 2078–2079. doi: 10.1093/bioinformatics/btp352

Luu, K., Bazin, E., and Blum, M. G. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. Mol. Ecol. Resour, 17, 67–77.

Madigan, D.J., Shipley, O. N., Carlisle, A. B., Dewar, H., Snodgrass, O. E., and Hussey, N. E. (2021). Isotopic tracers suggest limited trans-oceanic movements and regional residency in North Pacific blue sharks (Prionace glauca). Front Mar Sci, 8, 653606

Maxwell, S. M., Scales, K. L., Bograd, S. J., et al. (2019). Seasonal spatial segregation in blue sharks (Prionace glauca) by sex and size class in the Northeast Pacific Ocean. Diversity and Distributions, 25, 1304–1317.

MitoAnnotator. http://mitofish.aori.u-tokyo.ac.jp/annotation/input/

Musyl, M., Brill, R., Curran, D., and Fragoso, N. (2011). Postrelease survival, vertical and horizontal movements, and thermal habitats of five species of pelagic sharks in the central Pacific Ocean. Fish Bull, 109, 341–368.

Nakano, H., and Seki, M. P. (2003). Synopsis of biological data on the blue shark, Prionace glauca Linnaeus. Bulletin-Fisheries Research Agency Japan, 6, 18–55.

Nakano, H. and Stevens, J.D. (2008). The biology and ecology of the blue shark, Prionace glauca. In Camhi, M.D., Pikitch, E.K. and Babcock, E.A. (eds) Sharks of the open ocean: biology, fisheries and conservation. Oxford: Blackwell Publishing Ltd, pp. 140–151

Nielsen, E. E., Hemmer-Hansen, J., Foged, L. P., and Bekkevold, D. (2009). Population genomics of marine fishes: Identifying adaptive variation in space and time. Molecular Ecology, 18(15), 3128–3150.

Nikolic, N., Devloo, F., Bailleul, D., et al. (2022). Stepping up to Genome Scan Allows Stock Differentiation in the Worldwide Distributed Blue Shark Prionace Glauca. Mol. Ecol, 32, 1000–1019.

Peakall, R., and Smouse P.E. (2006) GenAlEx 6: Genetic Analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes, 6, 288– 295.

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research-an Update. Bioinformatics, 28(19), 2537–2539. doi: 10.1093/bioinformatics/bts460

Queiroz, N., Humphries, N. E., Noble, L. R., Santos, A. M., and Sims, D. W. (2012). Spatial dynamics and expanded vertical niche of blue sharks in oceanographic fonts reveal habitat targets for conservation. Plos One, 7(2), e32374.

R studio. https://posit.co/download/rstudio-desktop/

R Core Team. (2018). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. Agricultural Sciences, 9. https://www.R-project.org

Repeatmasker: <http://www.repeatmasker.org>

RepeatModeler: https://www.repeatmasker.org/RepeatModeler/

Sippel, T., Wraith, J., Kohin, S., et al. (2011). A Summary of Blue Shark (Prionace Glauca) and Shortfin Mako Shark (Isurus Oxyrinchus) Tagging Data Available from the North and Southwest Pacific Ocean. In Proceedings of the California: Working Document Submitted to the ISCShark Working Group Workshop, La Jolla, CA.

Sim, S. B., Corpuz, R. L., Simmonds, T. J., and Geib, S. M. (2022). Hifiadapterfilt, a memory efficient read processing pipeline, prevents occurrence of adapter sequence in PacBio HiFi reads and their negative impacts on genome assembly. BMC Genomics, 23(1), 157. doi:10.1186/s12864-022-08375-1.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA

polymorphism. Genetics, 123, 585-595

Taguchi, M., King, J. R., Wetklo, M., Withler, R. E., and Yokawa, K. (2015). Population Genetic Structure and Demographic History of Pacific Blue Sharks (Prionace Glauca) Inferred from Mitochondrial DNA Analysis. Marine and Freshwater Research, 66(3), 267–275. doi: 10.1071/MF14075

Vaux, F., Bohn, S., Hyde, J. R., and O'Malley, K. G. (2021). Adaptive markers distinguish North and South Pacific albacore amid low population differentiation. Evolutionary Applications, 14(5), 1343–1364. https://doi.org/10.1111/eva.132

Verissimo, A., Sampaio, Í., McDowell, J. R., et al. (2017). World without borders—genetic population structure of a highly migratory marine predator, the blue shark (Prionace glauca). Ecology and Evolution, 7, 4768–4781. https:// doi.org/10.1002/ece3.2987

Waples, R. S., Punt, A. E., and Cope, J. M. (2008). Integrating genetic data into management of marine resources: How can we do it better? Fish and Fisheries, 9(4), 423–449.

Whitlock, M. C., and K. E. Lotterhos. (2015). Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of FST. American Naturalist 186(suppl.), S24–S36.

Yu, Y., Ouyang, Y., and Yao, W. (2018). ShinyCircos: An R/Shiny Application for Interactive Creation of Circos Plot. Bioinformatics, 34(7), 1229–1231. doi: 10.1093/bioinformatics/btx763

**Tables**

Table 1. Summary of geographical area, sampling site, sampling year, and number of individuals for each analysis.



Table 2. Pairwise *F*ST value (below diagonal) and associated *P* values (above diagonal) among sampling sites based on the whole mitochondrial DNA sequence data set.



**Figures**

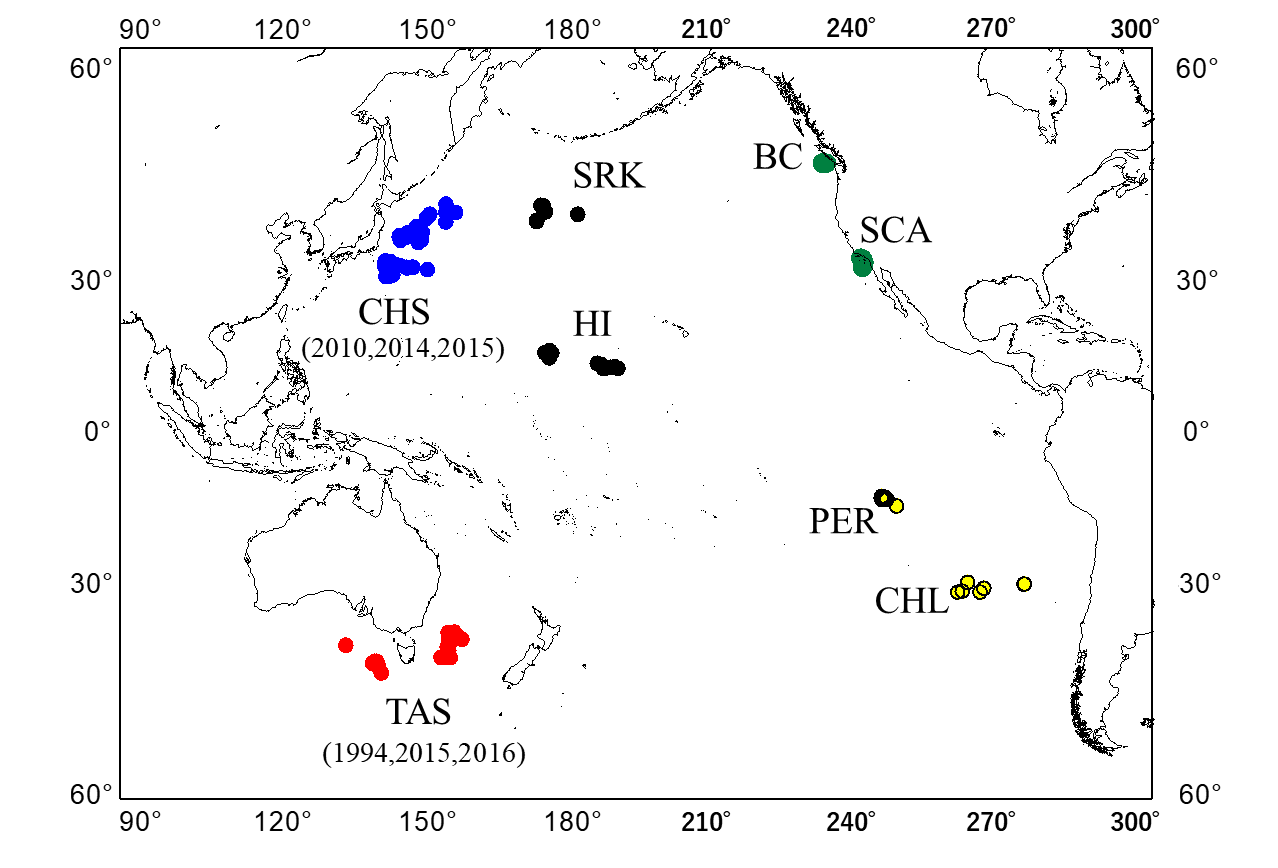


Fig.1 Sampling sites of blue shark, *Prionace glauca* in Pacific Ocean (288 individuals).

Geographical areas in North Pacific are off Choshi (CHS:2010, n=28; 2014, n=28; and 2015, n=24) in blue, off Sanriku (SRK:2011, n=24) and off Hawaii (HI:2011, n=24) in black, off British Columbia (BC:2010, n=20) and off Southern California (SCA:2010, n=24) in green. In South Pacific are off Tasmania (TAS:1994, n=22; 2015, n=24; 2016, n=28) in red, off Peru (PER:2019, n=29) and off Chile (CHL:2011, n=13) in yellow.

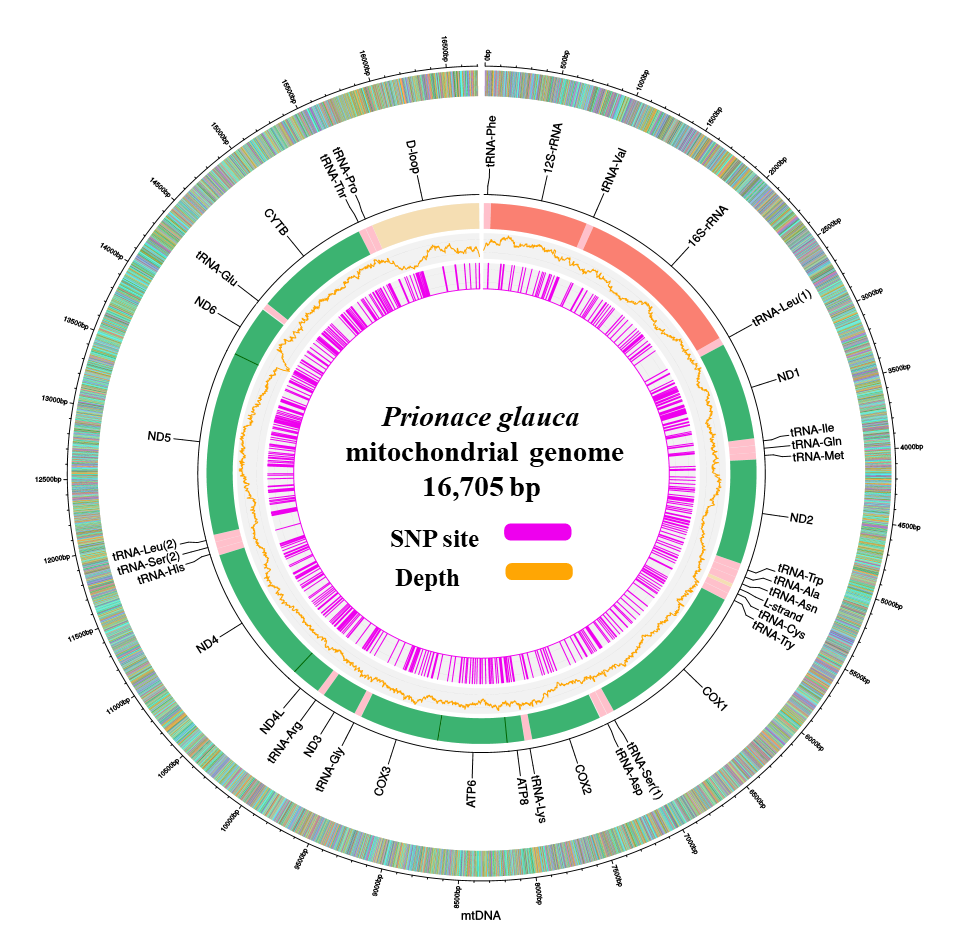
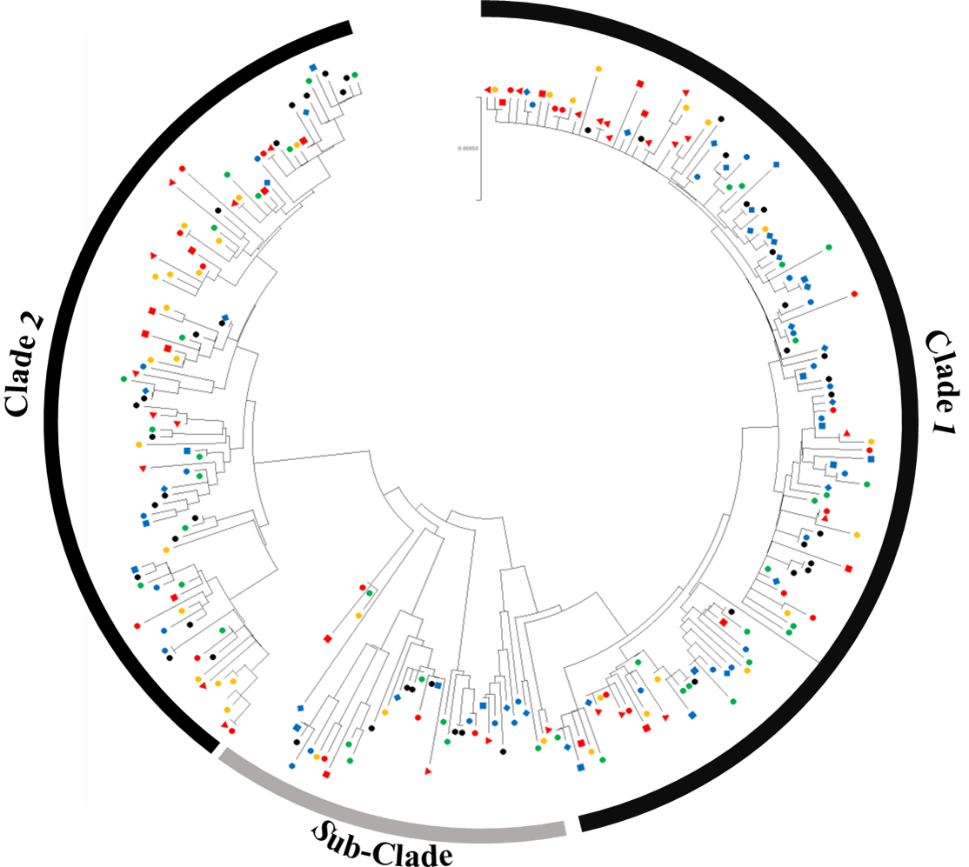


Fig.2 Circus plot of annotated Blue Shark mitochondrial gene region. ​Reference sequence, annotation, depth, and SNP information are shown from outer to inner circle.

(A)



（B）

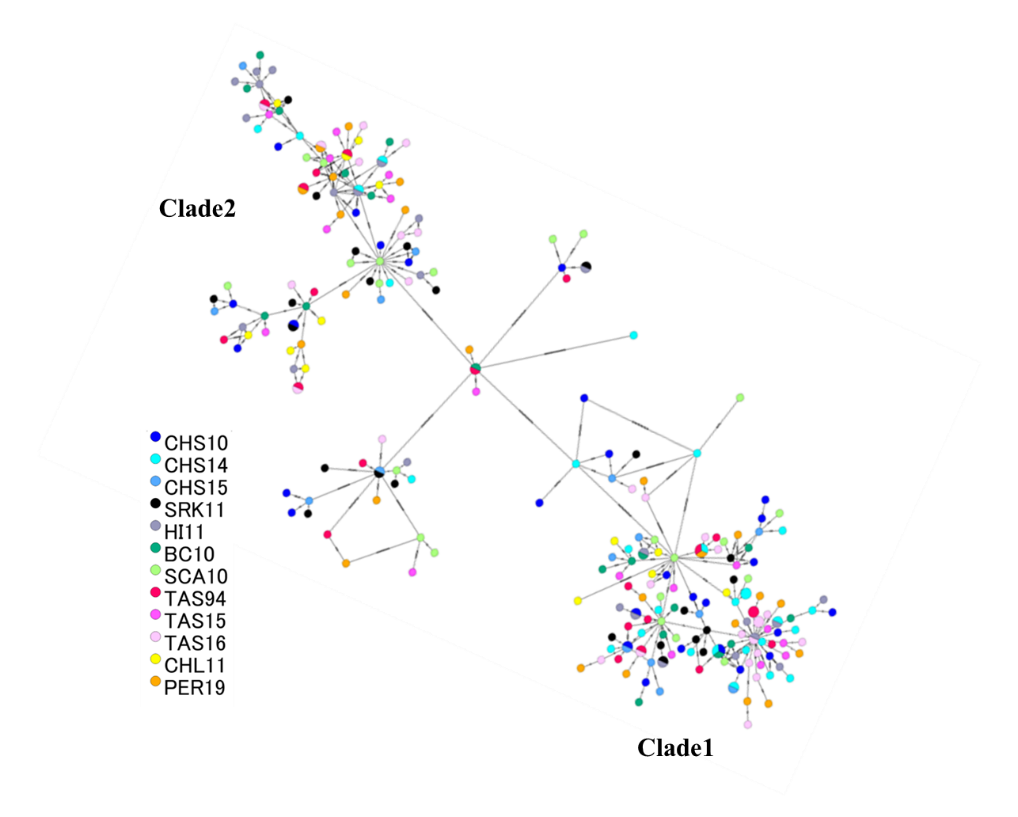


Fig.3 Phylogenetic tree (A) and haplotype network (B) of *Prionace glauca* from 264 individuals (238 haplotypes).

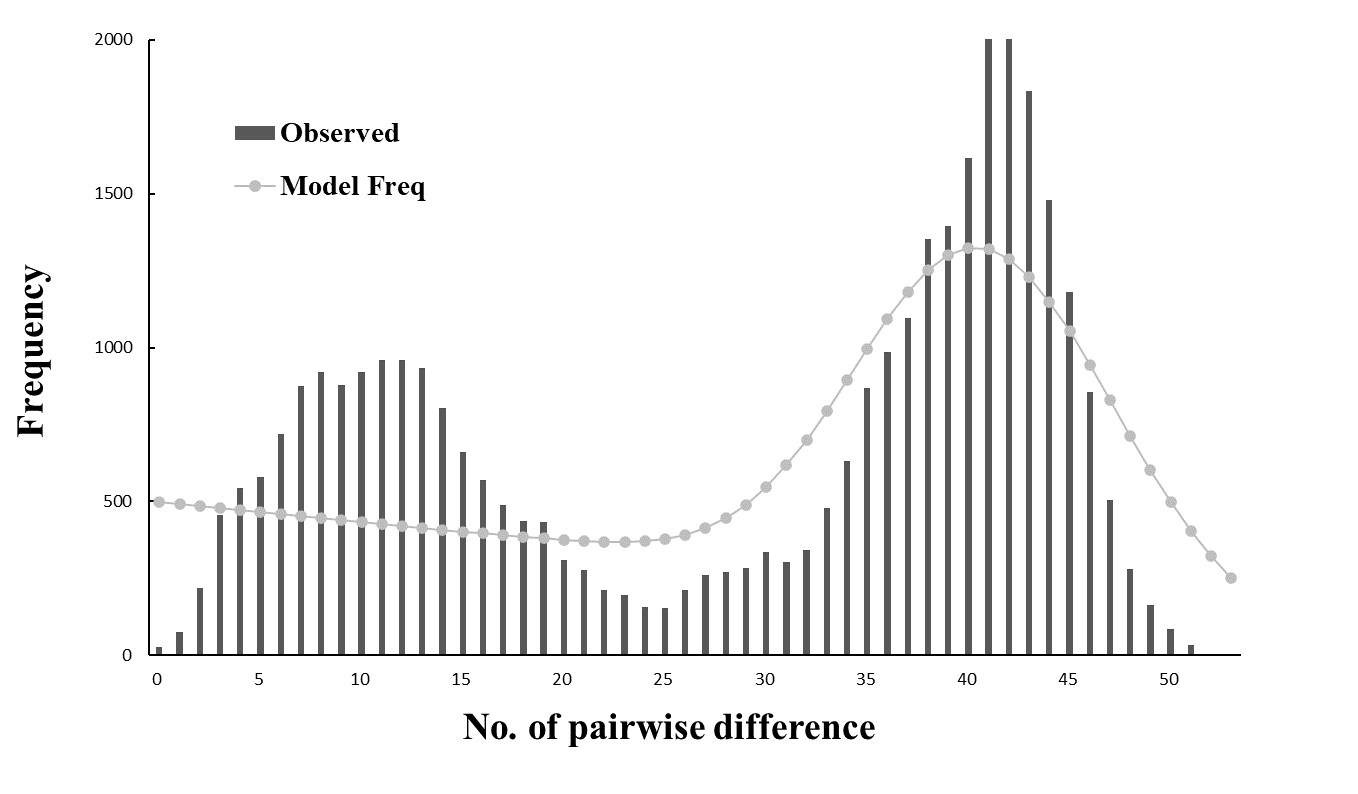
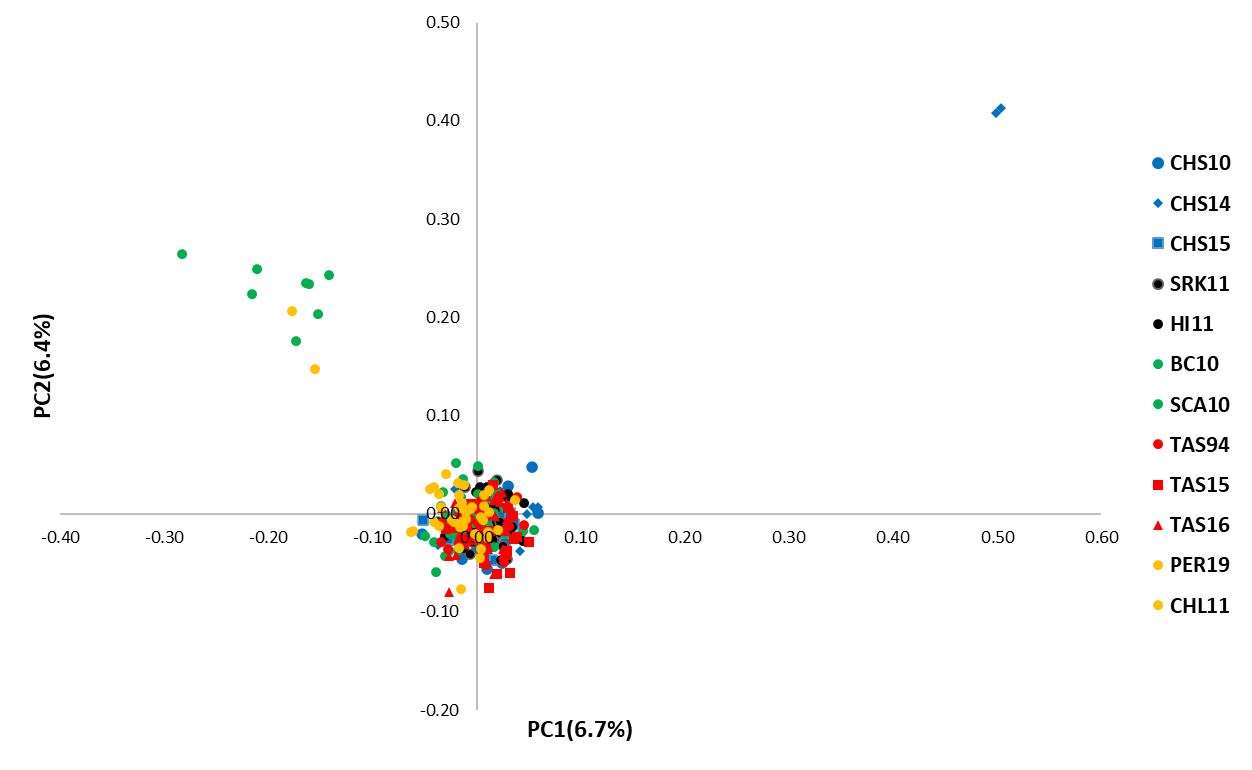


Fig.4 Mismatch distributions of *Prionace glauca* (264 individuals). Numbers of nucleotide substitutions are indicated along x-axis, and the frequency is indicated by y-axis.

(A)



(B)

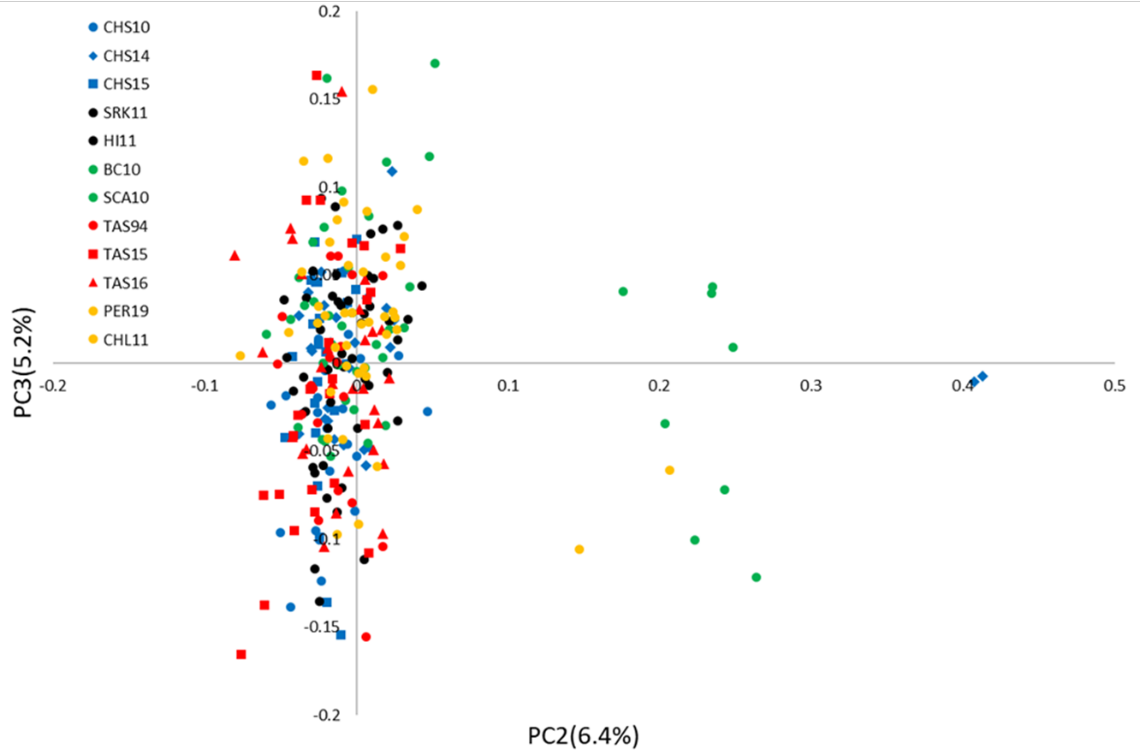
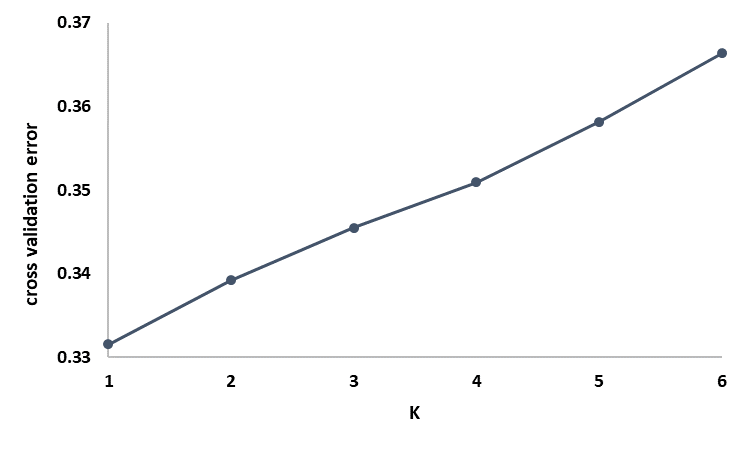


Fig.5 Principal component analysis (PCA) plot of first and second (A) and second and third (B) PC derived from 3,816 Gras-Di SNPs from 280 individuals.

(A)



.

(B)

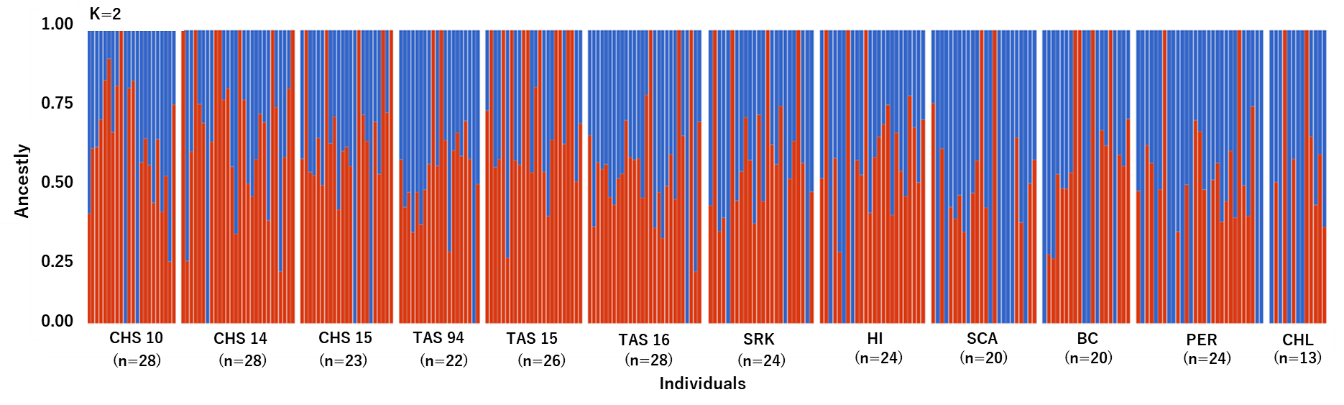


Fig.6 The cross-validation errors of the k values (A) and ADMIXTURE plots for K=2 where each bar represents a single individual, and the different colors reflect each of the K components (B).

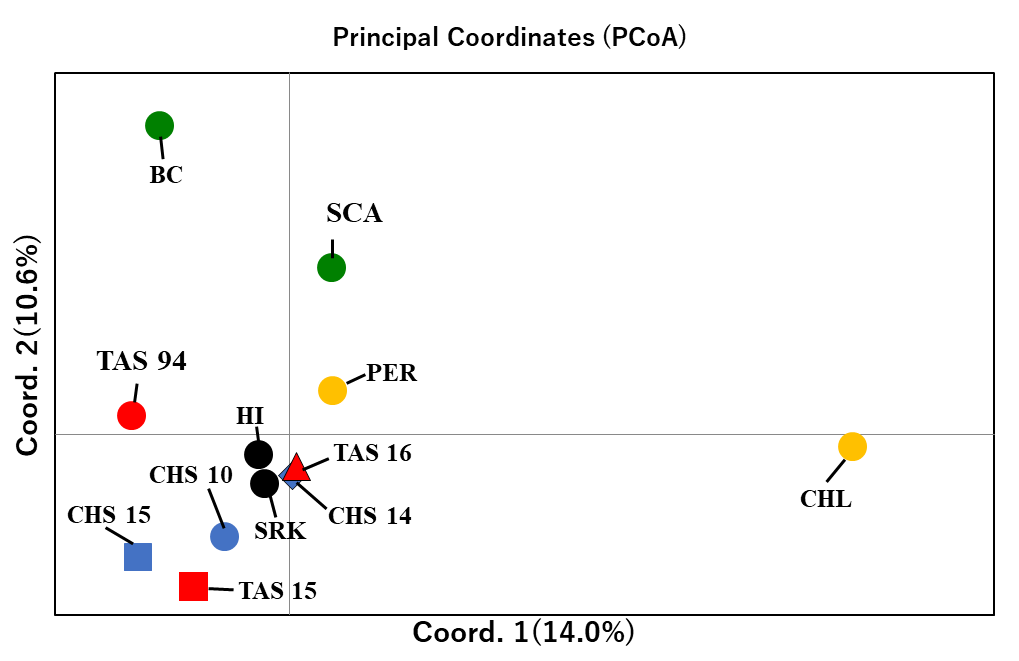


Fig.7 Principal coordinates analysis (PCoA) plots based on the pairwise-*F*ST values among sampling sites using Gras-Di data set.