國立臺灣大學生命科學院動物學研究所

碩士論文

Institute of Zoology

College of Life Science

National Taiwan University

Master Thesis

六斑二齒魨在西太平洋族群結構之研究

Population Structure of Long-spine Porcupinefish
(*Diodon holocanthus*) in the Western Pacific

蕭敦仁

Dun-Ren Hsiao

指導教授：邵廣昭 博士

施秀惠 博士

Advisors: Kwang-Tsao Shao, Ph.D.

Hsiu-Hui Shih, Ph.D.

中華民國 99 年 7 月

July, 2010

國立臺灣大學碩士學位論文
口試委員會審定書

六斑二齒魨在西太平洋族群結構之研究

Population Structure of Long-spine Porcupinefish (*Diodon holocanthus*) in the Western Pacific

本論文係蕭敦仁君（學號 R97B41024）在國立臺灣大學動物學研究所完成之碩士學位論文，於民國 99 年 7 月 19 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

王身壽 （指導教授）

陳章文 （指導教授）

王弘毅

陳春暉

系主任、所長 _陳俊宏_

Acknowledgement

中文摘要

　　六斑二齒魨（*Diodon holocanthus*）是二齒魨科（Family Diodontidae）的成員之中最為常見，數量也最為豐富的物種，在全球各大洋的近岸海域以環熱帶的模式分佈。由於六斑二齒魨缺乏明顯的成魚播遷行為與長距離移動能力，一般推測在其遼闊的分佈範圍裡，會因為基因交流受到距離限制而使得不同地區的族群之間產生遺傳組成之差異。本研究係以粒線體 DNA 中的控制區域片段（control region）和細胞色素 b（cytochrome b）基因來分析在西太平洋海域內，採集自台灣周圍、日本與兩者之間的六斑二齒魨個體，是否存在有能夠被偵測到的明顯族群結構。本研究以 Φ_{ST} 和 AMOVA 分析方法比較各區域之間遺傳組成之差異，建立序列之間的親緣關係樹狀圖和單倍體基因型（haplotype）之間的網狀結構，並進行變異分佈分析（mismatch distribution）與族群歷史的估算。研究結果顯示台灣與日本的族群之間並不存在顯著的遺傳差異，並且無法偵測到明顯的族群結構，顯示可能在不同地區的族群之間的基因交流並無受到距離限制，推測存在有促進基因交流的機制，例如甚長的漂浮仔稚魚期（pelagic larval duration），或是族群曾經歷過如冰河期等的歷史事件導致的遺傳組成均質化（genetic homogenization），使得族群分化的程度在此區域內並不顯著。

Abstract

The long-spine porcupinefish (*Diodon holocanthus*) is the most common and abundant species among all members of family Diodontidae, which has the circumtropical distribution pattern. As a marine fish with such wide geographic range but without known dispersal mechanism and the capability of long-distance migration, it is suspected that gene flows between populations of *D. holocanthus* should be limited, and thus in recognizable genetic differentiations among populations. We used the mitochondrial control region (D-loop) and cytochrome b (Cyt b) sequences to investigate whether the geographical structure exists among populations from Taiwan and Japan, within Western Pacific waters. The F-statistics analysis and AMOVA were applied for to detect genetic differences between populations. Phylogenetic trees and haplotype networks were constructed to elucidate the relationship among haplotypes. Demographic changes in the past were also addressed by analysis of mismatch distribution, as well as coalescence estimations. The result shows that the genetic difference between populations of different locations is insignificant. It indicates that gene flows between different populations are not limited by the geographical distances between sampling sites. Mechanisms may exist to enhance gene flows and thus prevent genetic differentiation, such as extended pelagic larval duration.

Contents

1. Introduction

1.1. Introduction to *Diodon holocanthus*

The Family Diodontidae, with common names such as porcupinefish, balloon fish or spiny puffer, is characterized by the scale-derived long spines covering most of the body surface, the reduced pelvic fins and the capability of inflation (Leis, 2001). In total, 6 genera, 19 species have been recognized all over the world. Among them, 3 genera, 6 species can be found around Taiwan waters (Shen et al., 1993), including the most common and abundant species, *Diodon holocanthus* L., the long-spine porcupinefish. *D. holocanthus* were found in warm seawaters and coasts around the world, from the Atlantic, the Indian Ocean, to the Pacific and the Oceanic islands, distributed in a circumtropical pattern (Figure 1). Within its native ranges, adult individuals reside solitarily in coastal waters, shallow reefs to open, and soft bottom (to at least 100 m depth). Their eggs, larvae and juveniles, however, are pelagic and can be caught offshore (Nishimura, 1960). Similar to many of its puffer relatives, *D. holocanthus* is a nocturnal hunter who possesses beak-like fused teeth, which act as crushing structure to be used to prey on hard-shelled benthic invertebrates, such as mollusks, crustaceans, and echinoderms (Leis, 2001).

As a species without much economical importance, the abundance and

commonness of *D. holocanthus* are the sole reason for theirs extensive appearances in literatures. However, the majority of previous studies focused on morphological descriptions, distributions and taxonomic status of the long-spine porcupinefish (Leis, 1978 and 2006a), and few others mentioned their physiological and mechanic mechanisms of inflation (Brainerd, 1994), food habitat (Pérez-España et al., 2005), toxicology (Hwang et al., 1992; Noguchi and Arakawa, 2008), parasitic infection (Kirtisinghe, 1934), and the cryopreservation of gonad and gamete (Chen, 2008). Biogeographic issues, in particular, the geographic genetic structure of this fish, have not been well investigated.

Despite their fishery importance is minimal (but still increasing, especially in Southern Taiwan or Pescadores), *D. holocanthus* sometimes appears in fish tank as an ornamental species, and adults can be kept in marine aquarium without much difficulty. As the result, many observations and recorded descriptions regarding its life cycle events, such as mating behavior and larval development, were made for the individuals that lived in artificial environment (Wolfsheimer, 1957; Leis; 1978; Sakamoto and Suzuki, 1978). While those observations may be biased from those in natural conditions, they still provide valuable information regarding the life cycle of *D. holocanthus*. There were also some observations based on natural populations or catching records of larvae or eggs (Nishimura, 1960). Combining both sources of observations from previous

2

studies, we can obtain preliminary knowledge regarding the life history of *D. holocanthus*.

The spawning event of diodontid in artificial tank was first described by Wolfsheimer (1957), but little descriptive information on eggs was mentioned. In addition, these eggs were failed to be fertilized, thus no further observation could be made on larval development (Wolfsheimer, 1957). After twenty years, two observational records of early life history of *D. holocanthus* were published independently by Leis (1978) and Sakamoto and Suzuki (1978). Both papers provided profound details about the spawning, embryogenesis, hatching, metamorphosis events of reared *D. holocanthus* individuals. Some observations, especially those made on morphological characters, were identical between the two papers. Both researches described that the fertilized egg was buoyant or pelagic, and the observed morphological characteristics of egg and larva were similar. But the measures regarding of morphometric characters, such as body size and length, of certain stage varied greatly. Except for days before hatching, which were around 4-5 days in both publications, larvae and juveniles of Leis showed a trend toward slower growth and longer development duration. Most notably, metamorphosis, the process that larvae grow and transform into juveniles, took place around 3 days after hatching in the paper of Sakamoto and Suzuki, but for Leis' record, metamorphosis occurred after 3 weeks. Nevertheless, combining the observations of both papers, the

pelagic larval duration (PLD) of *D. holocanthus* was at least 3 weeks, if we considered

metamorphosis signified the end of pelagic stage here.

Japanese scientists had published additional papers discussing reproduction events

and larval movement of *D. holocanthus*. Nishimura (1960) presented a table of more

than 70 collected "young-stage" (smaller than 100 mm in total length by his definition)

larvae of *D. holocanthus* in Western Pacific waters during the period of 1950s, and used

the information to infer the possible spawning site and putative larval migration routes.

Nishimura suggested that the migration of *D. holocanthus* might be classified into

"propagative migration", which would be easily affected by ocean currents, based on

the planktonic or pelagic nature of diodontid larvae. Fujita et al. (1997) showed the

annual reproductive cycle of *D. holocanthus* inferred by female GSI (gonad-somatic

index) and histological feature of ovaries. They deduced that there could be several

spawning events within the spawning season from May to June. They also speculated

that the possible spawning grounds of *D. holocanthus* might be not located near the

coast, since they could not collect females having ripe eggs along the coast during

spawning season.

It was supposed that *D. holocanthus* larvae would stay in offshore waters until they

grow to a fair size (100 mm by Nishimura's definition) before they recruit back to the

nearshore populations. The life history of early stages, after metamorphosis and before recruitment, is not clear. But events of mass recruitment, which by definition means recruiting with extraordinary huge number of individuals, were occasionally been spotted and recorded. For example, Debort and Nagelkerken (1997) published an observation on the mass recruitment event of juvenile *D. holocanthus* in Leeward Antilles. They hypothesized that *D. holocanthus* recruited primary in rare mass event and have extreme variation in settlement which similar to the case in another tetraodontiform, *Balistes vetula*, reported by Robertson (1988). Such mass recruitment event, followed by high mortality rate of recruits, was also spotted in Taiwan for several times (Hwang et al., 2009).

1.2. Dispersal ability and length of pelagic larval stage

For species that are highly residential during adult stage, dispersal by pelagic early-life stages determines patterns of population inter-connectivity. Pelagic larval stage can be found in many invertebrate and nearly all marine teleost fishes with demersal adult stage, and it is during this pelagic stage that the geographic scale of dispersal is determined for species with sessile or non-migratory adult stage (Leis, 2006b). Summarizing above previous studies, despite the fact that the migratory capability of adult *D. holocanthus* is limited by its poor swimmer ability, it is still

possible that *D. holocanthus* is capable of long-distance movement by the means of pelagic larval dispersal, although all evidences were indirect and this inference was tentative.

Pelagic egg has no mechanism for active movement, and its direction and speed of advection were determined by external forces only. As for larva, which was thought to have relatively limited ability for active locomotion, its dispersal is assumed to be heavily influenced by physical factors, such as wind, current or temperature (Gilg and Hilbish, 2003). No matter egg or larva, it can be assured that oceanographic or hydrodynamic conditions have strong influences on how and where pelagic early life may go. It is reasonable that under proper circumstances, such as strong mechanical forces provided by wind or surface oceanic current, buoyant particles can travel across a certain degree of range if they stay with the force long enough. If this is the case, one can easily deduce that the longer the length of pelagic duration, the longer the distance those buoyant particles can travel, in other term, better dispersal ability.

However, whether long pelagic larval duration does indeed have a positive correlation with dispersal ability or not is still in dispute, and evidence for low realized dispersal is accumulating from genetic and ecological studies for many marine species with a strong dispersal potential (Taylor and Hellberg, 2003; Johansson et al., 2008).

One problem is that the physical driving force must persist long enough to carry pelagic lifeforms away from their natal place. If the force ceased, or had varied considerably, the connections between populations by means of larval dispersal may not so solid. Another influencing factor is larval behavior. Contrasting to traditional knowledge, active behaviors of larvae, such as vertical migration, horizontal migration or even homing (retention or self-recruitment), have been recorded and documented for different marine taxa, as well as been proven to be a definite factor influencing the migrating route during pelagic larval stage, but such locomotion seems to be species specific (Leis, 2006b; Leis et al., 2007). And despite techniques such as otolith marking using tetracycline, tracking of larval movement in the open ocean is still limited in many other aspect. Thus Taylor and Hellberg (2003) concluded that, the simple assumption which extended pelagic stage will result in better dispersal ability is a faulty consideration at least for some taxa. Shanks (2009) revisited the relationship between pelagic duration and dispersal distance. While largely agreed upon the correlated relationship, he also emphasized that larval behavioral and species-specific life history traits play critical roles in dispersal distance.

As mentioned before, the pelagic larval duration of *D. holocanthus* is at least 3 weeks, with certain degrees of plasticity. Leis (1978) also commented that the length of pelagic stage is clearly must be measured in terms of month, and *D. holocanthus* had

been proven to be able to cross the eastern Pacific barrier (Lessios and Robertson, 2006). But in regions of western Pacific the population structure of *D. holocanthus* had never been investigated in detailed manner.

1.3.   Phylogeography and marine population structure

The field of biogeography deals with where organisms distribute in the world, and how the distribution was achieved (Avise, 2009). A recent branch of biogeography, phylogeography, focuses more on the principles and historical processes that responsible for the contemporary geographic distributions of genealogical lineages, especially within and among closed related species (Avise, 2009). Combing knowledge from population genetics and phylogenetics, those who studying phylogeography always infer insights from modern molecular data, in which the evidences for historic events hidden.

After the birth of neutral theory by Motoo Kimura, we know that most mutations which took place in the genome are relatively harmless, and thus can stay within the genome without being removed (Kimura, 1969). These neutral mutations, as well as mutated nucleotides that sit in non-functional or non-coding sequences, will "reside" in the genome for generations. If random mating was realized for a given population, mutations will be distributed across members of populations freely by maximized

intrapopulation gene flows. However, since neutral mutations occur in a random manner, if gene flows between groups of individuals were restricted by external factors, genetic homogenization is ought to be weakened and every group of a population accumulate mutations independently. This is how genetic differentiation comes about and these groups of population are on their way of divergence. By investigating population structure, the isolating or homogenizing forces underlie the formation of genetic differentiation can be revealed.

Collaborated with biological events in the past, such as dispersal or migration, extrinsic events along with their interactions contribute to the formation of modern population structure. Such extrinsic forces including geographical events, which limited where organisms lived, moved, and distributed, is the fundamental factor that hinders the movement of organisms and thus obstructs the gene flow. Geographical features that block the way gene flow goes may be insurmountable obstacles, such as water body for terrestrial animals, or simply a great amount of distance. On the other hand, climate changes such as glaciation, led to demographic changes. When glacier expanded during periods of global cooling, suitable habitats for a given species became compressed or fragmented and caused extensive reduction in population size. When glaciation ended, however, survivors which retreated to glacial refugia and lived there for many generations recolonized back to the habitable places once their ancestor lived, and

population size rose afterward. These demographic changes left evidences in genetic composition, such as population bottleneck or founder effect, which can be used for studying the history of populations (Templeton et al., 1995).

Comparing to the freshwater ichthyofauna, the marine fishes are thought to have a higher degree of genetic homogeneity, since biological barriers in the ocean are much more ambiguous and relatively inefficient (Palumbi, 1994). Oceanic trenches, for instance, are inoperative to species with pelagic phases. Also, the continuous water of marine environment represents a potential passage for migration, especially for nektonic species with good swimming performance, thus enhancing intermixing of genealogical lineages over the species range (Knutsen et al., 2003). Moreover, as partly described before, pelagic stages of marine organisms can potentially interconnect distant populations through dispersal on ocean currents, and species with such dispersing larvae should be genetically homogeneous over spatial scale (Taylor and Hellberg, 2003). But also as mentioned before, such interpretations must be made with cautions, for there are many other factors interplay within the relationship between pelagic stage and dispersal ability.

1.4.   Objective of research

The objective of this research is aimed to reveal the genetic structure of *D.*

holocanthus populations in the western Pacific waters, using mitochondrial sequences as applicable genetic markers. From results one can discuss that whether or not extraordinary reproductive attributes of this spiny organism, such as extended pelagic larval duration, did influence the connectivity between populations. Furthermore, by comparing the results with other sympatric marine animals, a better implication for the relationship of pelagic larval duration and dispersal ability can be inferred, and the role which oceanographic factors plays in facilitation of larval dispersal, can be investigated. Also from present-day genetic structure of *D. holocanthus*, as investigated in this study, can be used to draw inferences to reveal how the historical geographic and climatic changes shaping modern populations.
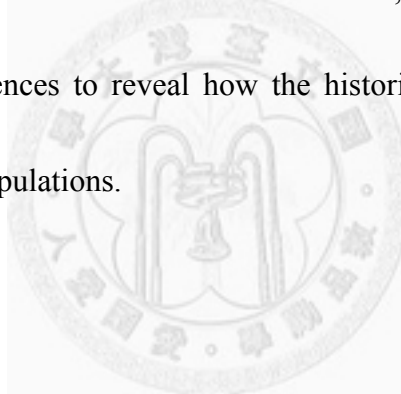
2. Materials and Methods

2.1. Sample collection, processing and preservation

2.1.1. Sampling range

Tissue samples of *Diodon holocanthus* used in this research were collected mainly from waters around Taiwan and Southern Japan. Other than that, some samples were captured from East China Sea by a Taiwanese R/V, *Ocean Researcher I*, during her cruise designated 905. The detailed information regarding to location and sample size of each sampling site was presented in Table 1, and distribution map of sampling sites were shown in Figure 2.

According to their geographical relationships, sampling sites were sorted into several locality groups, including JPN (Japan), RYK (Ryukyu), JEJ (Jeju Island), ECS (East China Sea), BEG (Beigan, Matsu), NTW (northern Taiwan waters), ETW (eastern Taiwan waters) and WTW (western Taiwan waters). Each group consisted of one or more sampling sites, and was used as the basic unit in some of the following analyses.

2.1.2. Collection, processing and preservation

Most samples were collected by hand-net, fixed net or angling. After capturing, samples were identified morphologically by several identification guides, including Fishes of Taiwan (Shen et al., 1993) and FAO species identification guide (Leis, 2001),

12
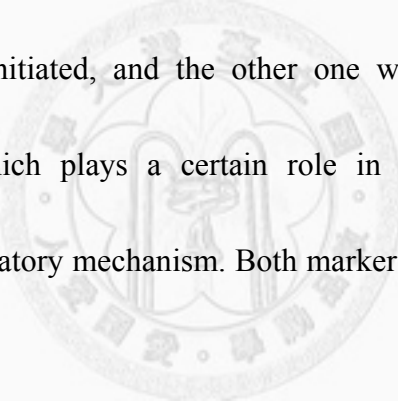
with measurements of weight and length (standard length) of some samples were recorded. Specific body part was cut off and preserved in 95% alcohol. Depends on the condition of sample, this body part could be dorsal fin, anal fin, pectoral fin or whole caudal peduncle. A few samples used in this research, especially those whose sample sites were not located near Taiwan, were collected by many foreign scientist, including: Naohide Nakayama and Hiromitsu Endo of Kochi University, Japan; Yohko Takata and Keiichi Matsuura of National Museum of Nature and Science, Japan; Choon Bok Song of Jeju National University, South Korea; H. J. Walker, Jr. of Scripps Institution of Oceanography, USA; Allan Connell of South African Data Centre for Oceanography, South Africa; Francesco Santini and Michael Alfaro of University of California, Los Angeles, USA; Sarah Samadi and Agnes Dettaï of Muséum national d'Histoire naturelle, France; and Tom Near and Kristen L. Kuhn, Yale University, USA. Eleven samples that are not *Diodon holocanthus* were collected together in order to used as outgroup in phylogenetic analyses. These samples are belonging to 4 different taxa within the Family Diodontidae, including *D. hystrix*, *D. liturosus*, *Chilomycterus reticulatus* and *Cyclichthys orbicularis*. Their sampling sites and details are not mentioned specifically.

Samples collected by R/V *Ocean Researcher I* were captured by trawling using zooplankton ring net with mesh size 1000 μm. Collected materials were fixed and preserved by 95% alcohol aboard, and eggs and larvae of *Diodon holocanthus* were

separated from other collected materials in laboratory under dissecting microscopes by

Mr. Han-Yang Chen and Ms. Mei-Yun Chu, as part of their own researches. After

separation, both eggs and larvae samples were preserved in 95% alcohol.
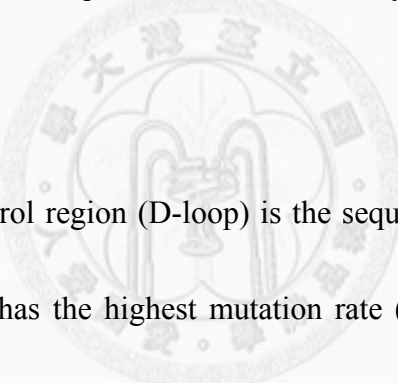
2.2.   Genetic markers

In this research, two genetic markers were utilized for revealing relationships

among samples from different locations. One was mitochondrial control region, also

known as D-loop or abbreviated as CR, a non-coding region at where replication of

mitochondrial genome is initiated, and the other one was the gene that codes for

Cytochrome b protein, which plays a certain role in electron transport chain, a

component of cellular respiratory mechanism. Both markers locate in the mitochondrial

genome.

The mitochondrial DNA (mtDNA), the nucleic acid that is carried by the organelle

which is accountable for cellular respiration, is a double-stranded circular molecule with

total length ranging approximately from 15 to 18 thousand base pairs. MtDNA has

many characteristics that differ from nuclear DNA, including haploidy, maternal

inheritance, rare heteroplasmy, and lack of introns, to name a few. Some of these

features make mtDNA an ideal genetic material that suits well in related to molecular

ecology, phylogenetics and population genetics. Also, even though protective or repair

mechanisms do exist in mitochondria, extraordinary high oxidative pressure still causes

the mutation rate of mtDNA five to ten times higher than nuclear DNA, which makes

mtDNA a practical choice for studying relationships at population level. Furthermore,

most cells, especially those with greater metabolic requirement, such as muscle cells,

have multiple mitochondria within a single cell, and full-length mitochondrial genome

of *Diodon holocanthus* had been published on the GenBank (Yamanoue et al., 2007),

which greatly reduced the efforts needed in the designing process of usable primer sets.

These characteristics make the amplification of mtDNA by PCR approach much more

feasible for studies.

The mitochondrial control region (D-loop) is the sequence among all segments of

mitochondrial genome that has the highest mutation rate (Meyer, 1993), thanks to its

non-coding attribute. As an appropriate genetic marker for addressing population-level

questions, D-loop was used widely in researches of different taxa, including fishes,

birds, mammals and many invertebrates. Cyt b, or the other hand, does not has a

mutation rate as extremely high as D-loop, but is still demonstrated by a variety of

studies to be an effective genetic markers when dealing with phylogenetic researches

that discuss lower-level relationships of different taxa.

Due to their convenience and efficiency when isolating haplotypes, and ideal

power of resolution as genetic markers for population level, both as mentioned above, D-loop and Cyt b were selected as the genetic markers in this study.

2.3. Acquiring target sequences

2.3.1. DNA extraction

The DNA molecules, which contain sequences that will be used as template in the following PCR amplification, were extracted and purified with the whole nuclear and mitochondrial genome, using a commercial DNA extraction kit (Genomic DNA Mini Kit, Geneaid Biotech Ltd.), which applies column with silica membrane for bounding to and separating DNA molecules from other cellular materials. Procedures described below followed the protocol provided by the kit's manufacturer, with minor modifications.

About 20 mg of tissue was cut off from sample which was preserved in 95% alcohol. The tissue was placed in 1.5 ml centrifuge tube, and was incubated at 55~60 °C in oven for 15 minutes in order to remove alcohol. After dealcoholization, 20 μl Proteinase K solution and 200 μl GT Buffer were added into the tube and then incubated at 60 °C for 1 hour. After all visible tissue had been dissolved, 200 μl GBT Buffer were added, and then incubated at 70 for 20 minutes. 200 μl absolute (99%) alcohol was added into the centrifuge tube after incubation was over, and mixed by vortex. The

well-mixed clear liquid with any precipitate then were transferred into a 2 ml GD

column with a silica membrane in the middle of the column. The GD column was

mounted on a collecting tube and then was centrifuged at 13200 rpm for 30 seconds,

using Centrifuge 5415 D (Eppendorf Inc.). Subsequently, 400 μl W1 Buffer were added

into the column, following by a second centrifugation, an addition of 600 μl Wash

Buffer and a third centrifugation with the same condition. A fourth centrifugation was

performed at 13200 rpm for 3 minutes in order to remove any remaining buffer. Before

the final centrifugation, all collected liquid was discarded and 70 μl Elution Buffer that

was preheated to 70 °C was added into the column. After 5 minutes for silica membrane

to fully absorb the extractant, the final centrifugation was performed at 13200 rpm for

30 seconds. After collecting, the purified DNA was stock in -20 °C for further

application.

2.3.2. Amplification of interested sequences

The polymerase chain reaction (PCR) was applied here for the amplification of

target sequences. Primers were designed based on whole mitochondrial genome of two

members of Family Diodontidae, *Diodon holocanthus* and *Chilomycterus reticulatus*

published on GenBank by Yamanoue et al. (2007), with accession number as AP009177

and AP009188, respectively. Two sets of primers were designed: DiL15535 (5'-GGT

CTT GTA AGC CGG ACG-3')/DiH123 (5'-GTC AGG ACC AAA CTT TTG TGC-3'),

whose target is the complete mitochondrial control region; and DiL14061 (5'-GAA

ACG GCT CCG CAG CCA AC-3')/DiH15551 (5'-CGG CTT ACA AGA CCG GCG

CT-3'), for amplification of nearly complete cytochrome b sequence. Figure 3 showed

the relationship between primer sets and the mitochondrial genome.

PCR reactant was prepared in 200 μl microcentrifuge tube, with final volume equal

to 25 μl. The PCR mix comprised of several well-mixed solutions, including: 2 μl

extracted DNA containing approximately 1 μg DNA, 2.5 μl 10× PCR Reaction Buffer

with 15 mM $MgCl_2$ (Genomics BioSci & Tech Co., Ltd.), 2 μl for each forward and

reverse primer (both 2.5 μM), 0.5 μl 40 mM dNTPs Mixtures (Genomics BioSci & Tech

Co., Ltd.), 0.5 μl 2.5U/μl *Taq* DNA polymerase (Genomics BioSci & Tech Co., Ltd.),

and 16 μl ddH_2O.

PCR reaction was performed using a DNA Engine Dyad Peltier Thermal Cycler

(Bio-Rad Laboratories, Inc). The program for amplification consists of an initial

denaturation stage (94 °C, 4 minutes), following by 35 cycles of denaturation (94 °C, 30

seconds), annealing (52 °C and 60 °C for D-loop and Cyt B, respectively, 30 seconds),

and elongation stages (72 °C, 1 minute), and a final extension stage (72 °C, 10 minutes).

PCR products were examined by agarose gel electrophoresis method. 5 μl PCR

product was loaded into the wall of 1% agarose gel after mixed with 1 μl 6× Loading Dye (Genomics BioSci & Tech Co., Ltd.). The gel for casting contained 0.05 μl/ml HealthView Nucleic Acid Stain (Genomics BioSci & Tech Co., Ltd.), as an alternative DNA stain to traditional ethidium bromide. The electrophoresis was performed under 100V for approximate 30~40 minutes, using MJ-105-S electrophoresis system (Major Science Co., Ltd.). Then the gel was inspected using gel documentation system UVCI-01-312 (Major Science Co., Ltd.), when gel was transilluminated by UV light, visible bands from valid PCR products can be detected.

Sequencing of validated PCR products was performed by a commercial company (Genomics BioSci & Tech Co., Ltd.), who utilizes Applied Biosystems 3730XL DNA Analyzer for DNA sequencing. PCR products of both markers (CR and Cyt b) were sequenced in both directions in order to increase accuracy and to eliminate artificial errors as possible.

2.4. Analyses of molecular data

2.4.1. Sequences management and alignment

Raw sequencing results were examined manually and individually, using Chromas v2.33 (Technelysium Pty. Ltd.), and those with substantial noise were removed from further analyses. D-loop and Cyt b sequences were compare to reference sequences,

namely *Diodon holocanthus* whole mitochondrial genome that was published on GenBank by Yamanoue et al. (2007), in order to check length and nucleotide position and to remove any artificial nucleotide substitution and insertion/deletion. Moreover, since Cyt b is a coding region, Cyt b sequences were checked to ensure that no stop codons exist along the length.

Sequences were managed using BioEdit 7.0.9.0 (Hall, 1999) and the conversions of sequence alignments between different file types were applied using online program ALTER (ALignment Transformation EnviRonment, Glez-Peña et al., 2010). Multiple sequence alignment was executed using MUSCLE method (MUltiple Sequence Comparison by Log-Expectation, Edgar, 2004a and 2004b) on an online application incorporated within the website of European Bioinformatics Institute. The results of alignment were checked and corrected manually.

2.4.2. Assessment of genetic diversity

Before any analyses were applied on acquired sequences, saturation tests were executed using DAMBE v5.1.1 (Xia and Xie, 2001) in order to ensure the markerss were suitable for following analyses. By pairwisely comparing estimated numbers of transition and transversion versus genetic distance between sequences, the degree of multiple substitutions of selected genetic markers can be revealed in the resulted plots.

If the degree of multiple substitutions was severe, as so called "saturated", phylogenetic information of this marker may be lost and should be discard for its lack of credibility.

DnaSP v5 (Librado and Rozas, 2009) was executed for computing various indices of genetic diversity including nucleotide diversity (π), haplotype number (h) and haplotype diversity. These indices of every individual location group and of whole population were calculated separately.

The haplotype number simply represents how many unique sequences exist in a given population. The haplotype diversity, another measure of haplotype uniqueness, was calculated with the formula described below:

$$h = n \, (1 - \Sigma f_i^2) / (n - 1)$$

where f_i is the haplotype frequency of the ith haplotype, and n is the sample size. The value of haplotype diversity ranges from its theoretical maximum 1 to minimum 0. If the haplotype diversity of a given population reaches 1, it means that every single sequence in this population is unique. On the other side, nucleotide diversity deals more with the difference of nucleotide composition, rather than haplotype, of a given population, thus acts as a measure of nucleotide polymorphism. The formula for calculating nucleotide diversity goes as following:

$$\pi = \Sigma f_i f_j \, \pi_{ij}$$

where f_i and f_j are the corresponding frequencies of the ith and jth sequences, π_{ij} is the proportion of nucleotide differences between ith and jth sequences, and n is the sample size.

2.4.3. Detection of population structure

2.4.3.1. F-statistics

F-statistics, also known as fixation index, is a statistical index for evaluating the degree of genetic fixation and heterozygosity. Its application is extended to estimate the degree of genetic differentiation between two given populations. If the comparison of two populations shows great and significant genetic differentiation, a population structure with a subdivided pattern can be inferred from F-statistics results. The software Arlequin 3.11 (Excoffier et al., 2005) was applied for calculating pairwise Φ_{ST} values, with 10000 permutations testing for significance.

2.4.3.2. Analysis of molecular variance, AMOVA

Other than direct estimation of genetic differentiation, genetic structure among populations can also be assessed by hierarchical analysis of molecular variance. AMOVA, which stands for analysis of molecular variance, is an analysis method that was developed by Excoffier et al. (1992) for estimating the source of molecular variance. As the proportions of molecular variance that come from three different scales (within

population, among populations within groups, and among groups) differs, the subdivision pattern of investigated species can be inferred from data.

The setting of analysis scenarios were as following: (1) all locality groups were dealt as separated groups; (2) considering that Kuroshio Current may play a certain role in dispersal mechanism of *D. holocanthus*, as mentioned before, locality groups were separated into two geographic groups: west, which includes JEJ, ECS, BEG and WTW; and east, which includes other locality groups, based on whether the locality group may be affected by Kuroshio Current or not. AMOVA was also conducted by the software Arlequin 3.11 (Excoffier et al., 2005), with 10000 permutations testing for significance.

2.4.4. Reconstruction of phylogenetic relationships

2.4.4.1. Phylogenetic trees

Before constructing phylograms, proper nucleotide substitution model that fits the data best must be determined in order to acquire convincible results. The software jModelTest 0.1.1 (Posada, 2008) was executed for performing likelihood tests of various models and searching the fittest one. Nucleotide substitution model with highest likelihood value was then applied in reconstruction of phylogenetic relationships. Maximum Likelihood was applied for reconstructing the relationship among all sampled individuals. Calculations were computed via an online service, RAxML

BlackBox (Stamatakis et al., 2008), using the best nucleotide substitution model suggested by jModeltest, as described above, as well as estimating the proportion of invariable sites of the sequences.

As mentioned beforehand, in order to facilitate the reconstruction of phylogenetic relationship among *D. holocanthus* sequences, 9 D-loop and 4 Cyt b sequences generated from 11 outgroup samples were applied in the analyses. Since there is no previous study that addressed the intra-family relationship of Family Diodontidae, which is remain unclear, the roots of constructed phylograms were set at the node between Genus *Diodon* and other genera (*Chilomycterus* and *Cyclichthys*).

2.4.4.2. Haplotype network

Similar to phylogenetic tree, the haplotype network can be used to depict the mutual relationship among samples. But in haplotype network, every molecular operational taxonomic unit (MOTU) represents a single haplotype, rather than single sequence, and frequencies of haplotypes are emphasized in the network plot. Therefore, when reconstructing phylogenetic relationships, haplotype network can be generated in order to provide additional information which focusing on ancestral-descendent relationship.

Median-joining parsimonious haplotype networks based on Cyt b marker was

constructed using the software NETWORK 4.5.1.6 (Fluxus Technology Ltd.).

Parameter ε was set as default (ε = 0). Additional two *D. holocanthus* samples, which

were sampled at two distant places, specifically Bali, Indonesia and Gulf of California,

Baja California Sur, Mexico, were used in the process of haplotype network

construction.

2.4.5. Demographic and coalescence estimations

2.4.5.1. Neutrality tests

Different neutrality tests, such as Tajima's D test or Fu and Li's D test, are designed

by various authors to investigate that whether a given population is biased form the

estimation under neutral theory or not. When the result of neutrality test of a given

population was biased from its neutral status, one can be sure that nucleotide

substitution events in the investigated population is not random, which may indicate that

a possible selective force is in effect, or this population is experiencing an event of

demographic changing, which may be subdivision or expansion. Thus, neutrality tests

can be applied for revealing the demographic history of target species. Both neutrality

tests, namely Tajima's D (Tajima. 1989) and Fu and Li's D (Fu and Li, 1993) were

conducted using the software DnaSP v5 (Librado and Rozas, 2009).

2.4.5.2. Mismatch distribution

The historical demographic pattern of *D. holocanthus* was also investigated by the analysis of mismatch distribution, which was performed by the software Arlequin 3.11 (Excoffier et al., 2005), on both markers separately. Simulated distributions were generated under sudden expansion model with 10000 bootstrap replicates. The overall validity of the estimated demography was assessed by the Harpending's Raggedness index (r, Harpending, 1994) and SSD (sum of squared deviation) between observed and simulated distributions and their corresponding P-values. Both indices were used to test the goodness-of-fit of the real data to model-simulated ones.

The analysis of mismatch distribution also calculated a critical parameter, τ (tau), which must be used in the coalescence estimation formula, described in the following section. Mismatch distribution also calculated θ_0 and θ_1, representing the effective female population size (N_{ef}) at the time of the last common ancestor and the current effective female population size.

2.4.5.3. Coalescence estimation

Methods based on coalescence theory provide another way for investigating the history of a given population. Theoretically, the genetic materials possessed by any individual of an extant population were inherited from a single common ancestor. By

26

tracing extant alleles of a marker back to a single ancestral copy, which was possessed by the most recent common ancestor (MRCA), the genealogy as well the history of extant population can be addressed.

The time since the most recent common ancestor was estimated with coalescence formula, based on the concept of molecular clock (Rogers and Harpending, 1992):

$$t = \tau / 2u = \tau / 4\mu k$$

where t is the time since coalescence, u is the mutation rate for the entire sequence per year, μ is the mutation rate per site and k is the sequence length. The parameter τ was calculated in the previous analysis of mismatch distribution. Since mutation rates for both markers of *D. holocanthus* used in this study had not been estimated before, the mutation rates used for calculation ($4-8\%$ Myr^{-1} for D-loop and $1-2\%$ Myr^{-1} for Cyt b, both in percentage of divergence between lineages) were based on the average divergence rates of several comparisons summarized by Lessios (2008).

3. Results

3.1. Genetic diversity

From all collected 185 *Diodon holocanthus* specimens, 174 whole-length D-loop and 162 nearly-complete Cyt b sequences were amplified and sequenced successfully, with the average lengths of 865 and 1120 base pairs respectively. The sequence AP009177, which was sequenced by Yamanoue et al. was used as a reference in the alignment and was included in all subsequent analyses, since its sampling site (Nobeoka City, Miyazaki Prefecture, Kyushu, Japan) is within the sampling range of this study (Yamanoue, per. comm.).

The plots of saturation tests (Figure 4) showed estimated numbers of transition and transversion against pairwise genetic distance, as well as corresponding best-fit curves, based on both markers. Except for the curve of transition of D-loop, other curves were mostly linear and didn't reach a plateau, which indicates certain amount of multiple substitutions. As for the data of transition of D-loop, even if the curve was not completely linear, it was far from reaching the saturation plateau. Hence the results could be interpreted as neither D-loop nor Cyt b sequence of *D. holocanthus* were saturated in substitution, and both transition and transversion could be used in the following analyses.

3.1.1. Genetic diversity of D-loop sequences

The alignment of 174 D-loop sequences showed many nucleotide substitutions and a small amount of indels, which were used as informative sites as well in the following analyses. 154 (17.72%) of all 869 positions were variable sites, within which 97 sites were parsimony informative and 43 were singleton variable positions.

Overall nucleotide diversity (π) of D-loop sequences was 0.02000. Among different locality groups, nucleotide diversity index varied from lowest 0.01867 (JEJ) to highest 0.02251 (RYK), as presented in Table 1. In total, 146 unique haplotypes were observed among 175 D-loop sequences, with the overall haplotype diversity (h) equaled to 0.997. Similar to nucleotide diversity, numbers of haplotype and haplotype diversity varied among different locality groups. The Northern Taiwan (NTW) had a highest haplotype number (67) of all observed groups, which was possibly affected by the large sample size of NTW. It is worth to note that most of all groups had haplotype diversity reached 1, the theoretical maximum of this index, with only one exception (NTW). This meant that every single sequence in that group differed from other sequences and presented a unique haplotype.

3.1.2. Genetic diversity of Cyt b sequences

The alignment of 148 nearly-complete Cyt b sequences had 1124 usable sites. No

indels were observed, and the level of nucleotide substitution was minimal. Within 86 (7.30% of 1124) variable sites, there were 44 parsimony informative sites and 42 singleton variable sites. As a protein-coding region, substitutions in Cyt b sequences may result in amino acid replacement. In 84 comparable sites, 11 were non-synonymous and 73 were synonymous substitutions. Of all variable sites, more than 79% (66) appeared in the third position of codon, and about 20% (17) resided in the first position. Only one variable site located at the second position. Varied changing rate of different codon positions, as presented here, was consistent with those reported previously (Bofkin and Goldman, 2007).

The nucleotide diversity of all Cyt b sequences was 0.00420, and it ranged from lowest 0.00387 (JPN) to highest 0.00497 (BEG) in different groups. Similar to the finding of D-loop, haplotype numbers and diversity varied among different groups. Overall haplotype diversity was 0.973, and haplotype diversity of groups ranged from 0.970 (NTW) to 1.000 (RYK). The most obvious difference between these two markers is that, while the number of usable sequences of Cyt b was slightly lower than that of D-loop, the difference of haplotype number was substantial. 73 haplotypes could be identified form 161 Cyt b analyzed sequences in total. Most of those shared haplotypes could be observed in many different locations, which may be distant to each other. In contrast to D-loop, the haplotype diversity was high, but nucleotide diversity of Cyt b

was lower than D-loop. This finding is reasonable that, since Cyt b codes for a functional protein, the variation of its nucleotide composition is restricted.

3.2. Population structure

3.2.1. F-statistics

Φ_{ST} values of pairwise comparison between groups of *D. holocanthus* populations, and corresponding *P* values was calculated and presented in Table 2 (D-loop) and Table 3 (Cyt b).

For both markers, all pairwise Φ_{ST} values are smaller than 0.05 or even less than 0, and only a few of them were statistically significant.

The Φ_{ST} values represented the degree of genetic differentiation between populations that under comparison, and the corresponding *P* values indicated that whether Φ_{ST} values were significantly larger than 0. When the value was larger than 0.05, the two comparing populations were considered to be genetically differentiated, and the degree of differentiation increases as the Φ_{ST} approaches to 1, the theoretical maximum of F-statistics. Occasionally, in some cases negative values were obtained due to the imprecision of the algorithm, and should be interpreted as equal to 0 (Foster et al., 2006; Jaramillo et al., 2001), which stands for no differentiation between population. Summarized from the results of F-statistics, sampled *D. holocanthus* populations in

western pacific water in this study presented little genetic differentiation. In other words, gene flows between populations were not limited, and both markers showed congruent results.

On the other hand, there were some pairwise comparisons that showed a trend toward slightly genetic differentiation between populations. Base on D-loop sequences, when comparing ECS to other groups, some of resulting Φ_{ST} were larger than 0.01 and reaching significances and reached statistical significance (ECS-NTW: $\Phi_{ST} = 0.02314$, $P = 0.03742$; ECS-ETW: $\Phi_{ST} = 0.04270$, $P = 0.03990$), although not a considerable or even noticeable degree of genetic differentiation. This finding of D-loop was not congruent to that of Cyt b, however. Despite the fact that Φ_{ST} values of ECS-JEJ, ECS-BEG and ECS-NTW pairing were the biggest ones and the only ones that larger than 0.01 in Table 3, none of them significantly larger than 0, thus could be interpreted as an evidence for differentiation. This is reasonable that since the effect of limited gene flow is easier to be found on sequences with higher mutation rate, D-loop in this case, than relatively slower Cyt b.

3.2.2. AMOVA

Table 4 and Table 5 showed the results of analysis of molecular variance, based on D-loop and Cyt b sequences respectively. For scenario 1, which considered all groups to

be independent, no significant genetic structures for *D. holocanthus* were observed, as

Φ_{ST} values for both markers were not significantly larger than 0, and almost all genetic

variation came from within population level (99.44% and 102.54%). A very little

fraction of variation came from among population within group level, with a relatively

small Φ_{SC} value. This was another result that stood as proof of insignificant and

negligible population structure.

However, the amount of percentage of variation that came from within group level

slightly changed when testing scenario was different. In scenario 2, locality groups were

separated into two parts, considering their geographic relationship with the Kuroshio

Current. Variation that came from among groups level was slightly higher than that of

scenario 1 for D-loop (from -1.68% to 0.81%) but not Cytb (from 0.54% to 0.65%),

whilst within population level was still the dominant variation factor. Again these results

were not consistent based on different markers, similar to the results of F-statistics.

3.3.   Phylogenetic relationship

3.3.1. Phylogenetic trees

As suggested by jModelTest, the best substitution model for D-loop of acquired *D.*

holocanthus sequences was GTR+I+Γ, and the best substitution model for Cyt b was the

same. Best substitution models were used in following tree-building processes. Two

phenograms, as presented in Figure 5 and Figure 6 representing the results based on

D-loop and Cyt b respectively, were constructed for revealing the phylogenetic

relationship among all sampled individuals.

As the phenogram based on D-loop depicted, all *D. holocanthus* samples clustered

in a single monophyly, with high supporting values (86). This clade was separated by

large distances from samples of other congener or confamilial taxa, but branches inside

this clade were not distinct and lacked of strong statistical support, though. Some

partitions were even polytomous. Furthermore, any barely observable partition inside

the ingroup clade did not match to geographical relationships among samples.

Similar to the result of D-loop, constructed phenogram based on Cyt b sequences

showed little differentiation among ingroup *D. holocanthus* individuals. Since the

differences between Cyt b sequences were smaller than D-loop, the ingroup clade of Cyt

b was even more shallow, and branching between samples within the monophyletic

ingroup lineage were even more polytomous and ambiguous.

3.3.2. Haplotype networks

A median-joining haplotype network which revealed the relationships among Cyt b

haplotypes was constructed and presented in Figure 7. The resulted haplotype network

presented a star-like pattern, with a central haplotype positioned in the most

intermediate place. This central haplotype linked to all peripheral ones, and was considered as the most ancestral haplotype. The central haplotype was not the most common one, however, and it was not shared by all populations.

The constructed haplotype network presented here, resembled phylogenetic trees in previous section in one way: the lack of distinct partition pattern. As mentioned before, the haplotype network of Cyt b sequences was star-like, and even if certain minor clades existed, they did not have distinguishable geographical structure. Those common haplotypes, including the central one, were shared by many groups which may be spatially distant to each other.

Likewise, samples from distant places, namely Bali and Gulf of California, did not form distinct and isolated lineages from other samples. Instead, the sample from Bali had a haplotype that was identical to samples collected near Taiwan or Japan; haplotype of sample from Gulf of California was unique, but there were merely two steps away only between it and the central haplotype.

3.4. Demographic estimation

3.4.1. Neutrality tests

Two different neutrality tests, scilicet Tajima's D test and Fu and Li's D test, were applied for examining that whether the observed population had shown a trend toward

bias from state of equilibrium. If the result showed a value that departed from 0, no matter positive or negative, it means that the evolutionary history of examined sequences are not neutral, and thus there might be some kind of mechanism that drives this deviation.

As Table 6 demonstrated, all testing resulted in negative values, regardless of which method was applied, which group or which marker was tested. Except for Tajima's D testing on all D-loop sequences, the results of tests that using all D-loop or all Cyt b data had negative values with statistical significance. For testing on each group separately, only few were significant.

3.4.2. Mismatch distribution

The distributions of pairwise mismatch between sequences were generated and presented in Figure 8 (D-loop) and Figure 9 (Cyt b). Both came with corresponding statistical indices, specifically SSD and Harpending's Raggedness index (r).

Both resulted mismatch distributions inferred from the two markers revealed unimodel distribution pattern, that is, only one dominant peak appeared in the distribution plot. Moreover, observed distributions of mismatch frequencies fitted the curves of simulated frequencies based on sudden expansion model pretty well. In the result of Cyt b (Figure 8), this was supported by significantly minimal values of both

indices. However, the result of D-loop was not supported by statistical testing, but the

values of indices were even smaller, which may suggest a trend toward good fittingness

of observed data to model.

3.4.3. Coalescence estimation

The coalescence parameters, including those calculated by analysis of mismatch

distribution such as τ, θ_0 and θ_1, and calculated coalescence time t, were presented in

Table 7. Since the values of θ_0 based on both markers were 0, the magnitude of

population growth from most recent common ancestor to modern population was

uncountable, but by comapring θ_0 and θ_1, population growth over this time still seems to

be enormous. The time when population expansion initiated was at about 549 to 220

kya (thousand years ago), within the period of Pleistocene, which ranged from 2.6 mya

(million years ago) to 12 kya.

4. Discussion

4.1. Genetic diversity of *D. holocanthus*

In this study two segments of mtDNA, D-loop (mitochondrial control region) and Cyt b, were applied for deciphering the population structure of *D. holocanthus*. Speak of genetic diversity, the two markers had similar pattern in terms of haplotype diversity. For D-loop, almost all sampling site had a haplotype diversity index as high as 1; even though many of them were largely affected by insufficient sampling, the overall haplotype diversity was still very close to 1. The haplotype diversity of observed Cyt b sequences was relatively high and close to 1 as well, but still slightly smaller than that of D-loop, despite the fact that sample size of Cyt b was smaller. On the other hand, the pattern of genetic diversity concerning nucleotide diversity was different for the two markers, as D-loop shows higher degree of nucleotide diversification, the average value of nucleotide diversity index of Cyt b was five times smaller than that of D-loop (% π = 2.01 for D-loop, 0.42 for Cyt b). D-loop had higher mean values for both haplotype and nucleotide diversity indices, especially the latter.. It is a reasonable observation, since D-loop is the segment within mtDNA that has highest mutation rate and the only one with non-coding attribute, more mutations can be accumulated in D-loop. On the contrast, the Cyt b is a protein-coding sequence, mutation within Cyt b must be somewhat restricted.

38

The combination of high haplotype diversity with intermediate to low nucleotide diversity was frequently observed in many demersal fishes, such as croaker *Larimichthys polyactis* (Xiao et al., 2009) or damselfish *Pomacentrus coelestis* (Liu et al., 2008), and may be the result of historical demographic decline or migration by founders, followed by sudden population expansion event (Grant and Bowen, 1998). Also, large haplotype diversity could be maintained within the population under conditions such as large population size, life history traits which facilitate rapid population expansion. Moreover, high mutation rates of selected genetic markers would contribute to high haplotype diversity.

4.2. Pattern of population structure of *D. holocanthus*

Results of F-statistics, whose values were nearly universally minimal or negative, clearly showed that genetic differentiations between locality groups were relatively indistinct. By the data based on two different mtDNA markers, population subdivision was insignificant, and little genetic differentiation could be detected. The same inference was drawn from the result of analysis of molecular variance (AMOVA, both scenarios), as almost all genetic variance came from within population level, which means genetic differences among groups did not significantly larger than difference among individuals of same group.

Marine fish generally show lower levels of genetic differentiation among geographical regions than freshwater fish, due to higher dispersal potential by means of planktonic egg, larva, or adult stages, coupled with the higher connectivity resulted from expanse of water body. Following sections discuss the two most possible causes that shaped observed population structure in detail.

4.2.1. Strengthened larval dispersal ability

By comparing calculated coalescent time with the period when barrier formed, previous study (Lessios and Robertson, 2006) provided evidence that *Diodon holocanthus* likely was able to cross the eastern Pacific barrier, a biogeographic barrier which is basically the expanse of deep water without islands, but the authors didn't discuss the possible mechanism responsible for such ability. As mentioned in the introduction section, the length of pelagic larval stage of *D. holocanthus* is at least 3 weeks, and may extend beyond a month, which is quite long among marine teleosts.

By observing the annual change of female GSI, Fujita et al. (1997) suggested that spawning season of *D. holocanthus* likely took place from late April to early June. Yun-Chi Liao keeps a record of number, weight and length of *D. holocanthus* individuals which impinged onto the protective mesh before the inlets of nuclear power plants of northern Taiwan (Liao, unpublished data). This eight-years-long record

presented that during the period from January to April, monthly average number of impinged *D. holocanthus* peaked, while length and weight per individual does not (Figure 10). Although harsh winter, from December to February, might have negative effects on weight of animals, but data of length showed that smaller body size should resulted from individuals of younger age, that is, juveniles. So this circumstance might indicate that the number of recruits in northern Taiwan waters maximizes during February to April, and these recruiting juveniles must be born in the previous year. If all these inferences were correct, that many pelagic larvae, if not all, stay in the open sea after being born and hatched for more than half a year, and, if carried by proper oceanographic force, could travel across a great distance. Another evidence for extended pelagic duration was inferred from otolith records (Ogden and Quinn, 1984), in which the ring counts of sagittae observed from "just-recruited" individuals exceeded 400. Ed Brothers, the researcher who made this observation, assumed that juveniles of *D. holocanthus* spent more than a year at somewhere before recruitment, based on the experience that many tropical fishes have daily microincrements on otolithes (E. Brothers, per. comm.). This provides another support for extended pelagic larval duration of *D. holocanthus*.

There are two dominant factors determining the direction and velocity of oceanic surface currents around Taiwan: East Asia monsoon system and the Kuroshio, whilst the

41

former being the primary force causing variation (Wyrtki, 1961). Kuroshio Current is well-known for its relatively high velocity, which may surpass 0.2 m/s (Liang et al., 2003, Figure 11). Under this circumstance, if eggs and larvae were carried by the Kuroshio Current, as speculated by Nishimura (1960), multiplying with the extra long pelagic duration, as discussed above, migration range should be considered in units of hundreds of kilometers or more. Such ability of migration, in absence of strong movement in adult stage, might still capable for dispersing individuals to and thus connecting gene pools of subpopulations.

4.2.2. Past demographic change

D. holocanthus seems to bear life history traits which capable of population expansion. Even though the detailed information regarding to life history, such as generation time, life span, or the age at first spawning, of *D. holocanthus*, were not scientifically documented, one may make conjectures regarding to life history traits of *D. holocanthus* based on known facts. A close relative of *D. holocanthus*, the ocean sunfish *Mola mola*, keeps the record for being the most fecund of all vertebrate, with a 137 cm female containing an estimated 300 million eggs (Pope et al., 2010). If this amazing trait of *M. mola* was shared by *D. holocanthus*, exponential population growth is highly likely to occur. Also, since the life expectancy of *M. mola* is longer than one

year, it may be iteroparous as well, which means it spawns multiple times in the entire

life. According to various unofficial observations by aquarists, the life span of *D. holocanthus* in aquarium ranges from 5 to 15 years, and might be iteroparous as well.

Above inferences are just assumptions, however, and might be incorrect, but the

commonness and abundance of extant *D. holocanthus* populations alone seem to be

enough for making the deduction that *D. holocanthus* is capable for rapid population

expansion.

The result of neutrality tests, analysis of mismatch distribution, and haplotype

network in this study, however, provided more direct evidence for possible past

expansion event. Both neutrality tests that applied here, namely, Tajima's D and Fu and

Li's D, showed significant negative values on both markers (Table 6). Negative value of

neutrality test signified the excess of low frequency polymorphism, which indicate the

presence of positively selective pressure or population expansion event. Support for

population expansion event was further strengthened by the result of a star-like

haplotype network (Figure 7) as well as unimodal mismatch distributions (Figure 8 and

9). Slatkin and Hudson (1991) pointed out that, for an exponentially growing population,

phylogeny will be nearly star-like and the distribution of pairwise differences (mismatch)

will be nearly a Poisson distribution, which corresponded to results of this research.

Recent population expansion was a common circumstance among many marine species,

and climate changes combining glaciation often underlay such demographic changes. More discussion regarding glaciation will be addressed in section 4.3.

4.2.3. Other facts concerning the observed population structure

Different scenarios of AMOVA (Table 4 and Table 5) showed a possible factor that might affect the formation of population structure: the Kuroshio Current. While the Kuroshio Current may play as a primary driving force that carries pelagic lifeforms of *D. holocanthus* across a great range, this strong current might impede dispersal with direction that is perpendicular to the current. However this effect, if does exist, was not obvious based on the data of this study, and more samples were required for addressing this assumption.

In a previous research Lessios and Robertson (2006) found *D. holocanthus* in the Eastern Pacific did have population structure with differentiation pattern. However, the geographical distance between the two comparing populations in that paper was more than 4000 kilometer, which was significantly larger than the sampling range of this study. Whether or not the population of *D. holocanthus* in the Western Pacific showed similar pattern to their eastern counterpart still requires more samples for this question to be resolved.

4.3. Glacial refugia and origin of extant population

When glaciation took place, mass mortality of marine animals was ought to be occur, due to the shrinkage of habitats caused by lowering of sea level. And postglacial expansion restored population size and habitat range after deglaciation, but molecular evidences of bottleneck and expansion event were written within the genome. Such glaciation and deglaciation cycle occurred multiple times during the history of Pleistocene, from 2.6 mya to 12 kya (Imbrie et al., 1992), and population expansion events we observed should occur after last glacial maximum (LGM), occurring from 110 kya to 10 kya. When last glacial maximum arrived, sea level lowered about 120 to 140 m, exposing much of suitable habitat which previously underwater, and population size of marine animal declined as consequence. Survival individuals retreated into locations that still habitable, which is so called "glacial refugia". After deglaciation, marine animals recolonized previously habitable location from glacial refugia, with population size grew in an exponential manner. Such events left some invisible traces in genetic composition of species, including evidences of sudden expansion which can be detected by analysis of mismatch distribution and neutrality tests, as mentioned before, and the most ancestral subpopulation, which can be revealed by reconstructing the phylogenetic relationship of extant individuals. The places where ancestral populations reside often indicate the presence of glacial refugia in the past.

4.4. Comparing to sympatric species

Population structure of many other marine animals in this area, that is, the western Pacific, had been investigated. Although different animals have diverse life history characteristics, by comparing them with *D. holocanthus* some abiotic factors or biotic properties that affecting population structure may be revealed. Two different patterns of population structure were frequently observed on marine animals of this region, specifically, the mixing pattern and the distinctly differentiated pattern.

Population structure of mixing pattern was often presented by little differentiation between geographical groups and a single inseparable clade comprising all sampled individuals, which were similar to the results of this study. Researches of sympatric fishes with this pattern including: two croakers *Nibea albiflora* (Han et al., 2008) and *Larimichthys polyactis* (Xiao et al., 2009), goby *Sicyopterus japonicus* (Leander, 2009), and Japanese eel *Anguilla japonica* (Han et al., 2010). Summarizing the results of these studies, two factors were referred as the key causes that shaped the undifferentiated genetic structure: migratory behavior at some stages, and prolonged pelagic larval stage. Migratory behavior and route of pelagic juvenile (leptocephalus) and adult of *Anguilla japonica* were well-studied and supposed to enhance the gene flows between populations in different locations (Han et al., 2010). Moreover, the migrating duration of

pelagic leptocephalus is estimated as 5 months. Similarly, *Nibea albiflora* has a pelagic

duration up to one month, and adult migrates between overwintering areas; both

contribute in homogenizing the genetic compositions in the region. *Sicyopterus*

japonicus also has a pelagic larval duration with the length from 5 to 6 months. The

length of pelagic duration for *Larimichthys polyactis* is unclear, however, but this

demersal species does spawn pelagic eggs, which might enhance its dispersal ability.

Even though *D. holocanthus* does not have migratory adult stage, its pelagic larval

duration is similarly long comparing to species described above, and therefore it is

reasonable to observe a universal population structure without local differentiation.

Population structure of distinct pattern was recognizable as certain genetic features

were observed, including: considerable genetic distance between lineages or

geographical groups, and distinguishable phylogenetic separations. These genetic

differentiation patterns, however, are not necessarily linked to geographical or temporal

relationships of organisms. Sympatric fish instances for such pattern includes: snapper

Lutjanus erythropterus (Zhang et al., 2006), damselfish *Pomacentrus coelestis* (Liu et

al., 2008), mackerel *Scomber australasicus* (Tzeng et al., 2009), and mullet *Mugil*

cephalus (Jamandre et al., 2009). The damselfish, *Pomacentrus coelestis*, spawns

demersal eggs in clutches. Demersal phase of egg may reduce the duration of pelagic

stage and thus distance this organism could travel by means of pelagic dispersal, and are

prone to be isolated by barriers in the seafloor, such as the Tokara gap, as Liu et al. (2008) suggested. Conversely the eggs of *D. holocanthus* are pelagic, and may have better ability to cross such submarine barriers. The other three examples, *Lutjanus erythropterus*, *Scomber australasicus* and *Mugil cephalus*, all showed distinct population structure but in a geographic scale that is different from this study. All of these three studies investigated populations of their own species in the western Pacific, including East China Sea (ECS) and South China Sea (SCS). They found that genetic distances between populations of ECS and SCS were considerable, but little differentiations were observable within ECS. It was assumed that the land bridge connecting Taiwan and Mainland China formed during glaciation blocked the gene flows between the two marginal seas. Actually those results are not inconsistent to this study, since *D. holocanthus* also presented an inseparable lineage in the ECS region. However, samples of *D. holocanthus* in the SCS were unable to be collected; therefore it is impossible to inspect that whether *D. holocanthus* have a similar structuring pattern to other fishes, until more samples, especially from other regions, were collected.

4.5. Conclusion

By investigated the population structure in the region of western Pacific, the fact that *Diodon holocanthus* is capable of long-distance migration, as indicated by the

existence of gene flows between populations, was revealed. Notwithstanding adults are not designed for active migratory movement, dispersal of *D. holocanthus* cross great geographical range is achieved by extended pelagic larval duration, in combination of local oceanographic features, such as the Kuroshio Current. Also, *D. holocanthus* was proven to experienced sudden population expansion after glaciation, a common historical events that many tropical fishes underwent.

5. References

Avise, J.C., 2009. Phylogeography: retrospect and prospect. *Biogeography* 36 (1): 3–15.

Bofkin, L., Goldman, N., 2007. Variation in Evolutionary Processes at Different Codon Positions. *Molecular Biology and Evolution* 24 (2): 513-521.

Brainerd, E.L., 1994. Pufferfish Inflation: Functional Morphology of Postcranial Structures in *Diodon holocanthus* (Tetraodontiformes). *Morphology* 220: 243-261.

Chen, Y.N., 2008. Cryopreservation of Male Gamate of Marine Vertebrate and Invertebrate – Studies in Porcupine Fish (*Diodon holocanthus*) and White Shrimp (*Litopenaeus vannamei*). Master thesis, Institute of Oceanography, National Taiwan University (in Chinese with English abstract).

Debrot, A.O., Nagelkerken, I., 1997. A rare mass recruitment of the balloonfish (*Diodon holocanthus* L.) in the Leeward Dutch Antilles, 1994. *Caribbean Journal of Science* 33 (3–4): 284–286.

Edgar R.C., 2004a. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatic*s 5: 113.

Edgar R.C., 2004b. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32 (5): 1792-1797.

Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics* 131 (2): 479-491.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin (version 3.0): An integrated

software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47-50.

Foster, C.B., Aswath, K., Stephen, J., Chanock, S.J., McKay, H.F., Peters, U., 2006. Polymorphism analysis of six selenoprotein genes: support for a selective sweep at the glutathione peroxidase 1 locus (3p21) in Asian populations. *BMC Genetics* 7: 56.

Fu, Y.X., Li, W.H., 1993. Statistical Tests of Neutrality of Mutations. *Genetics* 133 (3): 693-709

Fujita, T., Hamaura, W., Takemura, A., Takano, K., 1997. Histological Observatios of Annual Reproductive Cycle and Tidal Spawning Rhythm in the Female Porcupine Fish *Diodon holocanthus*. *Fisheries Science* 63 (5): 715-720.

Gilg, M.R., Hilbish, T.J., 2003. The geography of marine larval dispersal: coupling genetics with fine-scale physical oceanography. *Ecology* 84 (11): 2989-2998.

Glez-Peña, D., Gómez-Blanco, D., Reboiro-Jato, M., Fdez-Riverola, F., Posada, D., 2010. ALTER: program-oriented format conversion of DNA and protein alignments. *Nucleic Acids Research*, Web Server issue. ISSN: 0305-1048. http://dx.doi.org/10.1093/nar/gkq321

Grant, W.S., Bowen, B.W., 1998. Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. *Heredity* 89: 415-426.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41:

95-98.

Han, Y.S., Hung, C.L., Liao, Y.F., Tzeng, W.N., 2010. Population genetic structure of the Japanese eel *Anguilla japonica*: panmixia at spatial and temporal scales. *Marine Ecology Progress Series* 401: 221-232.

Han, Z.Q., Gao, T.X., Yanagimoto, T., Sakurai, Y., 2008. Genetic population structure of *Nibea albiflora* in Yellow Sea and East China Sea. *Fishery Science* 74: 544-552.

Harpending, H.C., 1994. Signature of ancient population growth in a low resolution mitochondrial DNA mismatch distribution. *Human Biology* 66 (4): 591-600.

Hsieh, H.J., Hsien, Y.L., Jeng, M.S., Tsai, W.S., Su, W.C., Chen, C.A., 2008. Tropical fishes killed by the cold. *Coral Reefs* 27: 599.

Hwang, D.F., Kao, C.Y., Yang, H.C., Jeng, S.S., Noguchi, T., Hashimoto, K., 1992. Toxicity of Puffer in Taiwan. *Nippon Suisan Gakkaishi* 58 (8), 1541-1547.

Hwang, J.S., Shao, K.T., Cheng, I.J., Hu, J.H., Huh, C.A., Lo, W.T., Fang, T.H., 2009. Ecological survey on coastal waters of nuclear power plants in Northern Taiwan: Final yearly report of 2009. National Taiwan Ocean University (in Chinese).

Imbrie J., Boyle, E.A., Clemens, S.C., Duffy, A., Howard, W.R., Kukla, G., Kutzbach, J., Martinson, D.G., McIntyre, A., Mix, A.C., Molfino, B., Morley, J.J., Peterson, L.C., Pisias, N.G., Prell, W.L., Raymo, M.E., Shackleton, N.J., Toggweiler, J.R., 1992. On the Structure and Origin of Major Glaciation Cycles 1. Linear Responses to Milankovitch Forcing. *Paleoceanography* 7 (6): 701–738.

Jamandre, B.W., Durand, J.D., Tzeng, W.N., 2009 Phylogeography of the flathead

mullet Mugil cephalus in the north-west Pacific as inferred from the mtDNA control region. *Fish Biology* 75 (2): 393-407.

Jaramillo, C., Montaña, M.F., Castro, L.R., Vallejo, G.A., Guhl, F., 2001. Differentiation and Genetic Analysis of *Rhodnius prolixus* and *Rhodnius colombiensis* by rDNA and RAPD Amplification. *Memórias do Instituto Oswaldo Cruz* 96 (3): 1043-1048.

Johansson, M.L., Banks, M.A., Glunt, H.M., Hassel-Finnegan, H.M., Buonaccorsi, V.P., 2008. Influence of habitat discontinuity, geographical distance, and oceanography on fine-scale population genetic structure of copper rockfish (*Sebastes caurinus*). *Molecular Ecology* 17 (13): 3051-3061.

Kaschner, K., Ready, J.S., Agbayani, E., Rius, J., Kesner-Reyes, K., Eastwood, P.D., South, A.B., Kullander, S.O., Rees, T., Close, C.H., Watson, R., Pauly, D., Froese, R. 2008. AquaMaps: Predicted range maps for aquatic species. World wide web electronic publication, www.aquamaps.org, Version 10/2008.

Kimura, M., 1969. The number of heterozygous nucleotide sites maintained in a finite population due to the steady flux of mutations. *Genetics* 61: 893-903.

Kirtisinghe, P., 1934. Parasitic Infection of Porcupine Fish. *Nature* 133: 142.

Knutsen, H., Jorde, P.E., Andr, C., Stenseth, N.C., 2003. Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. *Molecular Ecology* 12: 385-394.

Leander, N.J.S., 2009. Population genetic structure of the amphidromous goby *Sicyopterus japonicus* in the northwestern pacific. Master thesis, Institute of Fishery Science, National Taiwan University (In English with Chinese abstract).

Leis, J.M., 1978. Systematics and Zoogeography of the Porcupinefishes (*Diodon*, Diodontidae, Tetraodontiformes), with Comments on Eggs and Larval Development. *Fishery Bulletin* 76 (1): 535-567.

Leis, J.M., 2001. Tetraodontiformes: Diodontidae. In Carpenter, K.E., Niem, V.H., eds. FAO species identification guide for fishery purposes. The living marine resources of the Western Central Pacific. Volume 6: Bony fishes part 4 (Labridae to Latimeriidae), estuarine crocodiles, sea turtles, sea snakes and marine mammals. Rome, FAO. 2069-2790.

Leis, J.M., 2006a. Nomenclature and distribution of the species of the porcupinefish family Diodontidae (Pisces, Teleostei). *Memoirs of Museum Victoria* 63 (1): 77-90.

Leis, J.M., 2006b. Are Larvae of Demersal Fishes Plankton or Nekton? *Advances in Marine Biology* 51: 57-141.

Leis, J.M., Wright, K.J., Johnson, R.N., 2007. Behaviour that influences dispersal and connectivity in small, young larvae of a reef fish. *Marine Biology* 153: 103-117.

Lessios, H.A., Robertson, D.R., 2006. Crossing the impassable: genetic connections in 20 reef fishes across the eastern Pacific barrier. *Proceedings of the Royal Society B: Biological Sciences* 237: 2201-2208.

Lessios, H.A., 2008. The Great American Schism: Divergence of Marine Organisms After the Rise of the Central American Isthmus. *Annual Review of Ecology, Evolution, and Systematics* 39: 63-91.

Liang, W.D., Tang, T.Y., Yang, Y.J., Ko, M.T., Chuang, W.S., 2003. Upper-ocean currents around Taiwan. *Deep Sea Research Part II: Topical Studies in*

Oceanography 50 (6-7): 1085-1105.

Librado, P., Rozas, J., 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.

Liu, S.Y.V., Kokita, T., Dai, C.F., 2008. Population genetic structure of the neon damselfish (*Pomacentrus coelestis*) in the northwestern Pacific Ocean. *Marine Biology* 154: 745-753.

Meyer, A., 1993. Evolution of mitochondrial DNA in fishes. In Hochachka, P.W., Mommsen, P., eds. *Biochemistry and molecular biology of fishes*. Vol 2. Elsevier Press, The Netherlands.

Nishimura, S., 1960. Some aspects of the natural history of porcupine puffers migrating to the Japanese waters: I. spawning and migration. *Japanese Journal of Ecology* 10 (1): 6-11 (in Japanese with English abstract).

Noguchi, T., Arakawa, O., 2008. Tetrodotoxin - Distribution and Accumulation in Aquatic Organisms, and Cases of Human Intoxication. *Marine Drugs* 6 (2): 220-242.

Odgen, J.C., Quinn, T.P., 1984. Migration in coral reef fishes: Ecological significance and orientation mechanisms. In McCleave, J.D., Arnold, G.P., Dodson, J.J., Neill, W.H., eds. *Mechanisms of migration in fishes*, Plenum Press, New York. 293-308.

Palumbi, S.R., 1994. Genetic Divergence, Reproductive Isolation, and Marine Speciation. *Annual Review of Ecology and Systematics* 25: 547-572.

Pérez-España, H., Saucedo-Lozano, M., Raymundo-Huizar, A.R., 2005. Trophic

ecology of demersal fishes from the Pacific shelf off central Mexico. *Bulletin of Marine Biology* 77 (1): 19-31.

Pope, E.C., Hays, G.C., Thys, T.M., Doyle, T.K., Sims, D.W., Queiroz, N., Hobson, V.J., Kubicek, L., Houghton, J.D.R., 2010. The biology and ecology of the ocean sunfish Mola mola: a review of current knowledge and future research perspectives. *Reviews in Fish Biology and Fisheries*, Corrected Proof, Available online DOI: 10.1007/s11160-009-9155-9.

Posada, D., 2008. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution* 25: 1253-1256.

Robertson, D.R., 1988. Extreme variation in settlement of the Caribbean triggerfish Balistes vetula in Panama. *Copeia* 1988 (3): 698–703.

Rogers A.R., Harpending, H.C., 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9 (3): 552-569.

Sakamoto, T., Suzuki, K., 1978. Spawning Behavior andEarly Life History of the Porcupine Puffer, *Diodon holocanthus*, in Aquaria. *Japanese Journal of Ichthyology* 24 (4): 261-270 (in Japanese with English abstract).

Shanks, A.L., 2009. Pelagic Larval Duration and Dispersal Distance Revisited. *Biological Bulletin* 216: 373-385.

Shen, S.C., Lee, S.C., Shao, K.T., Mok, H.K., Chen, C.H., Chen, C.T., 1993. Fishes of Taiwan. Department of Zoology, National Taiwan University, Taipei (in Chinese).

Slatkin, M., Hudson, R., 1991. Pairwise Comparisons of Mitochondrial DNA Sequences in Stable and Exponentially Growing Populations. *Genetics* 129: 555-562.

Stamatakis, A., Hoover, P., Rougemont, J., 2008. A Rapid Bootstrap Algorithm for the RAxML Web-Servers. *Systematic Biology* 75 (5): 758-771.

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

Taylor, M.S., Hellberg, M.E., 2003. Genetic Evidence for Local Retention of Pelagic Larvae in a Caribbean Reef Fish. *Science* 299: 107-109.

Templeton, A.R., Routman, E., Philips, C.A., 1995. Separating population structure from population history: a cladistic analysis the of the Geographical Distribution of Mitochondrial DNA Haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. *Genetics* 140: 767-782.

Tzeng, C.H., Chen, C.S., Tang, P.C., Chiu, T.S., 2009. Microsatellite and mitochondrial haplotype differentiation in blue mackerel (*Scomber australasicus*) from the western North Pacific. *ICES Journal of Marine Science*, available online DOI: 10.1093/icesjms/fsp120.

Wolfsheimer, G., 1957. A spawning of porcupine puffers. *The Aquarium*, Philadelphia, 26 (9): 288-290.

Wyrtki, K., 1961. Physical oceanography of the southeast Asia waters. *Scripps Institute of Oceanography NAGA Report*, 2. Scientific Results of Marine Investigations of the South China Sea and the Gulf of Thailand 1959-1961.

Xia, X., Xie. Z., 2001. DAMBE: Data analysis in molecular biology and evolution. *Journal of Heredity* 92: 371-373.

Xiao, Y., Zhang, Y., Gao T., Yanagimoto, T., Yabe, M., Sakurai, Y., 2009. Genetic diversity in the mtDNA control region and population structure in the small yellow croaker *Larimichthys polyactis. Environmental Biology of Fishes* 85: 303-314.

Yamanoue, Y., Miya, M., Matsuura, K., Yagishita, N., Mabuchi, K., Sakai, H., Katoh, M., Nishida, M., 2007. Phylogenetic position of tetraodontiform fishes within the higher teleosts: Bayesian inferences based on 44 whole mitochondrial genome sequences. *Molecular Phylogenetics and Evolution* 45 (1): 89-101.

Yamanoue, Y., Miya, M., Matsuura, K., Katoh, M., Sakai, H., Nishida, M., 2008. A new perspective on phylogeny and evolution of tetraodontiform fishes (Pisces: Acanthopterygii) based on whole mitochondrial genome sequences: Basal ecological diversification? *BMC Evolutionary Biology* 8: 212.

Zhang, J., Cai, Z., Huang, L., 2006. Population genetic structure of crimson snapper *Lutjanus erythropterus* in East Asia, revealed by analysis of the mitochondrial control region. *ICES Journal of Marine Science* 63: 693-704.

Table 1. Sampling details of *Diodon holocanthus* used in this study, with sample size (N), number of available sequences (n), haplotype number (n_h), haplotype diversity (h) and nucleotide diversity (π) of each locality group.

Locality group (abbreviation)	Sampling site (abbreviation)	N	D-loop				Cyt b			
			n	n_h	h	π	n	n_h	h	π
	JPN Overall	14	14	14	1.000	0.02243	14	13	0.989	0.00387
Japan (JPN)	Ashizuri Port, Tosa-Shimizu City, Kochi, Shikoku, Japan (AZP)	1	1	1	1.000	−	1	1	1.000	−
	Iburi Fishing Port, Tosa-Shimizu City, Kochi, Shikoku, Japan (IBR)	8	8	8	1.000	0.02150	8	8	1.000	0.00435
	Nakanoshima, Susaki City, Kochi, Shikoku, Japan (NNS)	1	1	1	1.000	−	1	1	1.000	−
	Okinoshima, Sukumo City, Kochi, Shikoku, Japan (ONS)	2	2	2	1.000	0.01736	2	2	1.000	0.00267
	Saga, Kuroshio Town, Kochi, Shikoku, Japan (SAG)	1	1	1	1.000	−	1	1	1.000	−
	Nobeoka City, Miyazaki, Kyushu, Japan (MZK)	1	1	1	1.000	−	1	1	1.000	−
Ryukyu (RYK)	RYK Overall	4	3	3	1.000	0.02251	4	4	1.000	0.00437
	Amami−oshima, Okinawa, Ryukyu, Japan (AMO)	1	0	−	−	−	1	1	1.000	−
	Kumejima, Okinawa, Ryukyu, Japan (KMJ)	3	3	3	1.000	0.02251	3	3	1.000	0.00415

Table 1, continued.

| Locality group (abbreviation)	Sampling site (abbreviation)	n	D-loop				Cyt b			
			n	n_h	h	π	n	n_h	h	π
Jeju Is., S. Korea (JEJ)	JEJ Overall	15	10	10	1.000	0.01867	13	11	0.974	0.00430
East China Sea (ECS)	ECS Overall	23	19	19	1.000	0.01903	20	18	0.989	0.00391
	Site 16: 124.12°E, 29.00°N	1	1	1	1.000	−	1	1	1.000	−
	Site 31/32: 123.06−124.09°E, 28.27−27.88°N	22	18	18	1.000	0.01882	19	18	0.994	0.00403
Beigan, Matsu (BEG)	BEG Overall	15	14	14	1.000	0.02010	15	14	0.990	0.00497
Northen Taiwan (NTW)	NTW Overall	75	75	67	0.996	0.02001	64	38	0.970	0.00422
	Nuclear Power Plant 1 Inlet (NP1)	17	17	16	0.993	0.01887	17	13	0.963	0.00423
	Nuclear Power Plant 2 Inlet (NP2)	19	19	17	0.988	0.02198	16	10	0.917	0.00380
	Nuclear Power Plant 4 Inlet (NP4)	39	39	37	0.997	0.02099	31	26	0.989	0.00443
Eastern Taiwan (ETW)	ETW Overall	16	16	16	1.000	0.02069	10	9	0.978	0.00467
	Hualien Harbor, Hualian (HUL)	6	6	6	1.000	0.01988	4	4	1.000	0.00489
	Chengkung, Taitung (CHK)	10	10	10	1.000	0.02170	6	6	1.000	0.00474
Western Taiwan (WTW)	WTW Overall	24	24	24	1.000	0.01920	21	16	0.971	0.00408
	Fongkuei, Penghu (FOK)	5	5	5	1.000	0.01676	5	5	1.000	0.00374
	Magong, Penghu (MAG)	10	10	10	1.000	0.01894	7	7	1.000	0.00322
	Liuqiu, Pingtung (LCH)	9	9	9	1.000	0.02313	9	8	0.972	0.00509
Total		186	175	146	0.997	0.02000	161	73	0.973	0.00420

Table 2. Pairwise comparison Φ_{ST} values between groups (below diagonal) and corresponding *P* values (above diagonal), inferred from mitochondrial control region (D-loop) of *Diodon holocanthus*.

	JPN	RYK	JEJ	ECS	BEG	NTW	ETW	WTW
JPN	−	0.92624	0.93723	0.28571	0.83625	0.89684	0.86387	0.72458
RYK	−0.08153	−	0.86873	0.52698	0.87823	0.85952	0.86081	0.72884
JEJ	−0.03930	−0.08239	−	0.43511	0.62509	0.72003	0.83873	0.60935
ECS	0.00861	−0.01639	−0.00056	−	0.29017	0.03742	0.03990	0.27423
BEG	−0.02344	−0.07024	−0.01312	0.00858	−	0.73913	0.53450	0.96822
NTW	−0.01595	−0.05449	−0.01318	0.02314*	−0.01043	−	0.50559	0.44679
ETW	−0.02577	−0.07139	−0.02986	0.04270*	−0.0066	−0.00278	−	0.27294
WTW	−0.01328	−0.03915	−0.01033	0.00687	−0.02800	−0.00072	0.00896	−
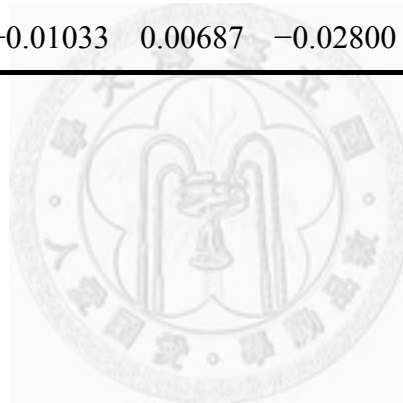
Table 3. Pairwise comparison Φ_{ST} values between groups (below diagonal) and corresponding *P* values (above diagonal), inferred from cytochrome b of *Diodon holocanthus*.

	JPN	RYK	JEJ	ECS	BEG	NTW	ETW	WTW
JPN	−	0.35096	0.90941	0.55697	0.57034	0.90466	0.99109	0.97782
RYK	0.00786	−	0.59153	0.40204	0.80636	0.73666	0.93654	0.54283
JEJ	−0.03037	−0.01752	−	0.15157	0.33205	0.83932	0.91961	0.65964
ECS	−0.00688	0.00797	0.02161	−	0.23176	0.08811	0.50817	0.53450
BEG	−0.00727	−0.04431	0.00761	0.01203	−	0.68845	0.89407	0.77844
NTW	−0.01705	−0.03310	−0.01469	0.01509	−0.00809	−	0.97980	0.90506
ETW	−0.04902	−0.05920	−0.03235	−0.00546	−0.03127	−0.03150	−	0.98564
WTW	−0.03068	−0.01211	−0.00976	−0.00484	−0.01610	−0.01245	−0.04456	−
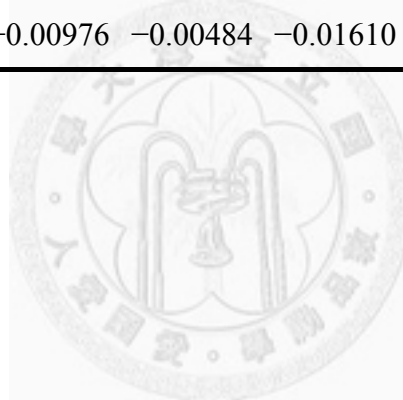
Table 4. Hierarchical analysis of molecular variance (AMOVA) based on D-loop sequences. Above: scenario 1, all groups were separated; below: scenario 2, grouping according to the relationship to the Kuroshio Current.

Source of variation	Degrees of freedom	Variance component	Percentage of variation	Fixation Index	P value
Among groups	7	−0.14752 Va	−1.68	$\Phi_{CT} = -0.01675$	0.88119
Among population within group	11	0.19703 Vb	2.24	$\Phi_{SC} = 0.02200$	0.11356
Within population	156	8.75726 Vc	99.44	$\Phi_{ST} = 0.00562$	0.18307
Among groups	1	0.04738 Va	0.54	$\Phi_{CT} = 0.00535$	0.13208
Among population within group	17	0.04658 Vb	0.53	$\Phi_{SC} = 0.00529$	0.28109
Within population	156	8.75726 Vc	98.94	$\Phi_{ST} = 0.01062$	0.17832

Table 5. Hierarchical analysis of molecular variance (AMOVA) based on Cyt b sequences. Above: scenario 1; below: scenario 2. Analysis scenarios setting were identical to that of D-loop.

Source of variation	Degrees of freedom	Variance component	Percentage of variation	Fixation Index	*P* value
Among groups	7	0.01904 Va	0.81	$\Phi_{CT} = 0.00808$	0.67752
Among population within group	12	−0.07878 Vb	−3.35	$\Phi_{SC} = -0.03372$	0.95238
Within population	141	2.41491 Vc	102.54	$\Phi_{ST} = -0.02537$	0.98416
Among groups	1	0.00344 Va	0.15	$\Phi_{CT} = 0.00146$	0.64752
Among population within group	17	−0.06097 Vb	−2.59	$\Phi_{SC} = -0.02590$	0.98119
Within population	141	2.41491 Vc	102.44	$\Phi_{ST} = -0.02441$	0.98248

Table 6. The result of two statistical neutrality test of *Diodon holocanthus*, using D-loop and Cyt b as genetic markers. Neutrality tests for group RYK were not executed because of insufficient sample size.

Locality groups	Tajima's D		Fu and Li's D	
	D-loop	Cyt b	D-loop	Cyt b
JPN	−0.66095	−1.68184	−0.31781	−1.45830
	P > 0.10	0.10 > P > 0.05	P > 0.10	P > 0.10
RYK	−	−0.82943	−	−0.82943
		P > 0.10		P > 0.10
JEJ	−0.59832	−1.73690	−0.52516	−1.68022
	P > 0.10	0.10 > P > 0.05	P > 0.10	P > 0.10
ECS	−0.89596	−1.64186	−0.94538	−1.83875
	P > 0.10	P > 0.10	P > 0.10	P > 0.10
BEG	−0.62816	−1.47512	−0.41082	−1.00554
	P > 0.10	P > 0.10	P > 0.10	P > 0.10
NTW	−1.25946	−1.99082*	−2.16297	−2.31937
	P > 0.10	P < 0.05	P > 0.10	0.10 > P > 0.05
ETW	−0.45440	−1.81382*	−0.70280	−1.93480
	P > 0.10	P < 0.05	P > 0.10	0.10 > P > 0.05
WTW	−0.78873	−1.81563*	−0.88348	−1.65190
	P > 0.10	P < 0.05	P > 0.10	P > 0.10
Total	−1.30339	−2.17225**	−2.20321*	−4.37791*
	P > 0.10	P < 0.01	P < 0.05	P < 0.05

Table 7. Estimated coalescence time and parameters of *D. holocanthus*, based on two different markers. Unit for *u* was percentage of divergence between lineages, and unit for *t* was year.

Coalescence	τ	θ_0	θ_1	u	Population age (t)
D-loop	18.973	0.000	416.250	2–4%	274000–549000
Cyt b	4.951	0.000	916.250	1–2%	220000–439000
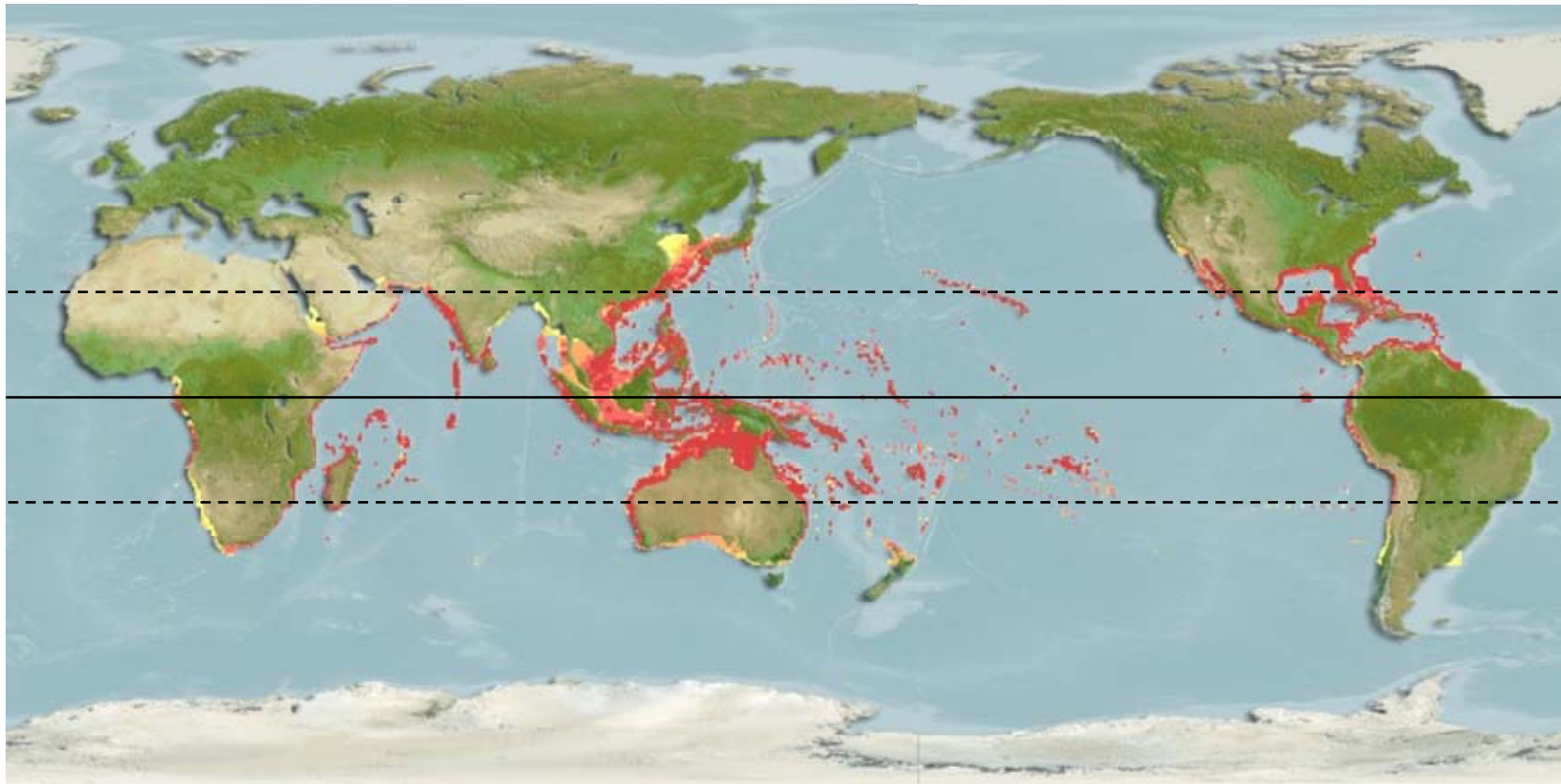
Figure 1. The native distribution range of *Diodon holocanthus*. Modified from AquaMaps (Kaschner et al., 2008). Solid line: the Equator; dash lines: the Northern and the Southern tropics.
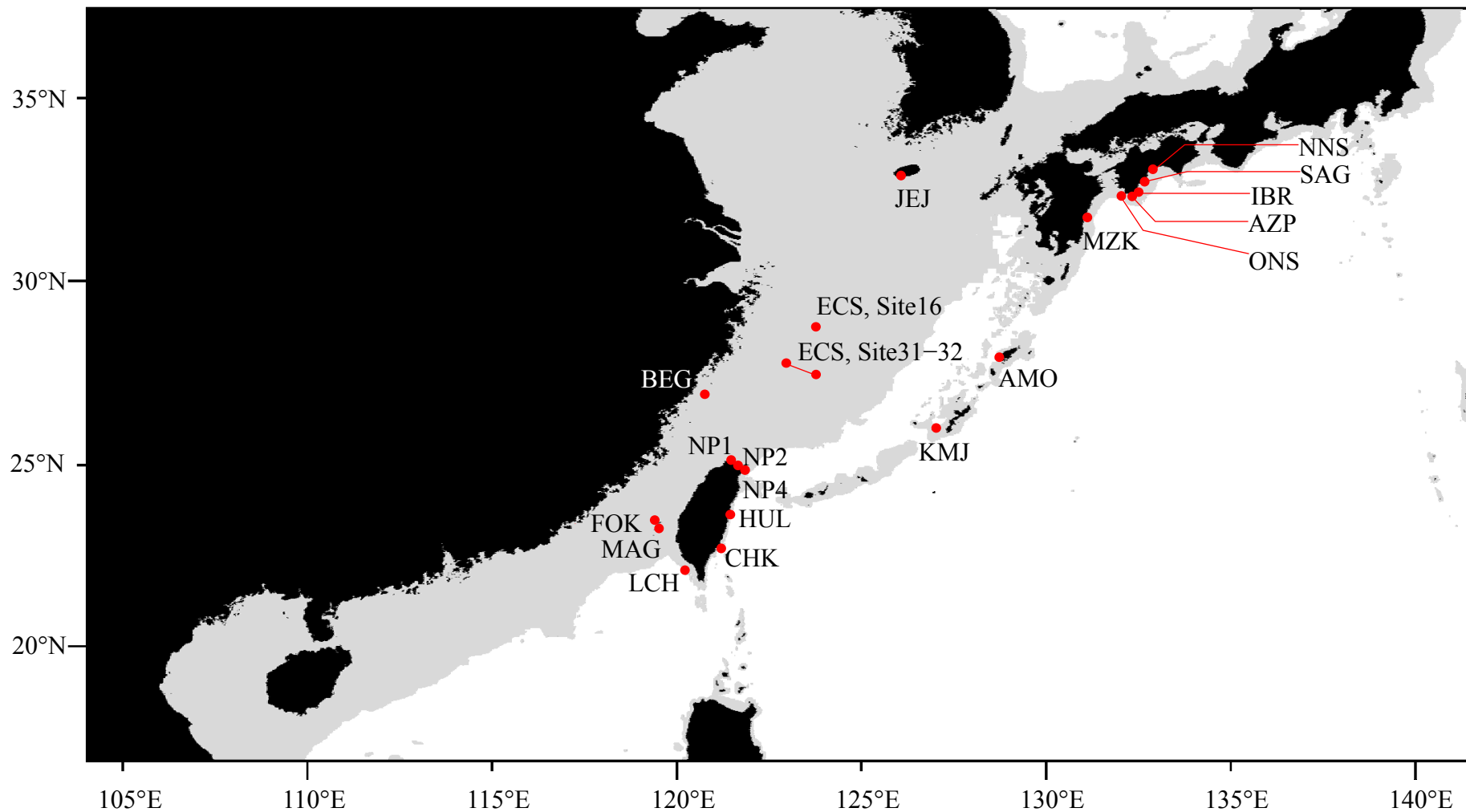
Figure 2. The sampling sites of *Diodon holocanthus*, around Western Pacific waters, in this study. Shaded area presented continental shelf with depth smaller than approximately 120 m. For complete name of abbreviations of sampling sites, seek table 1.
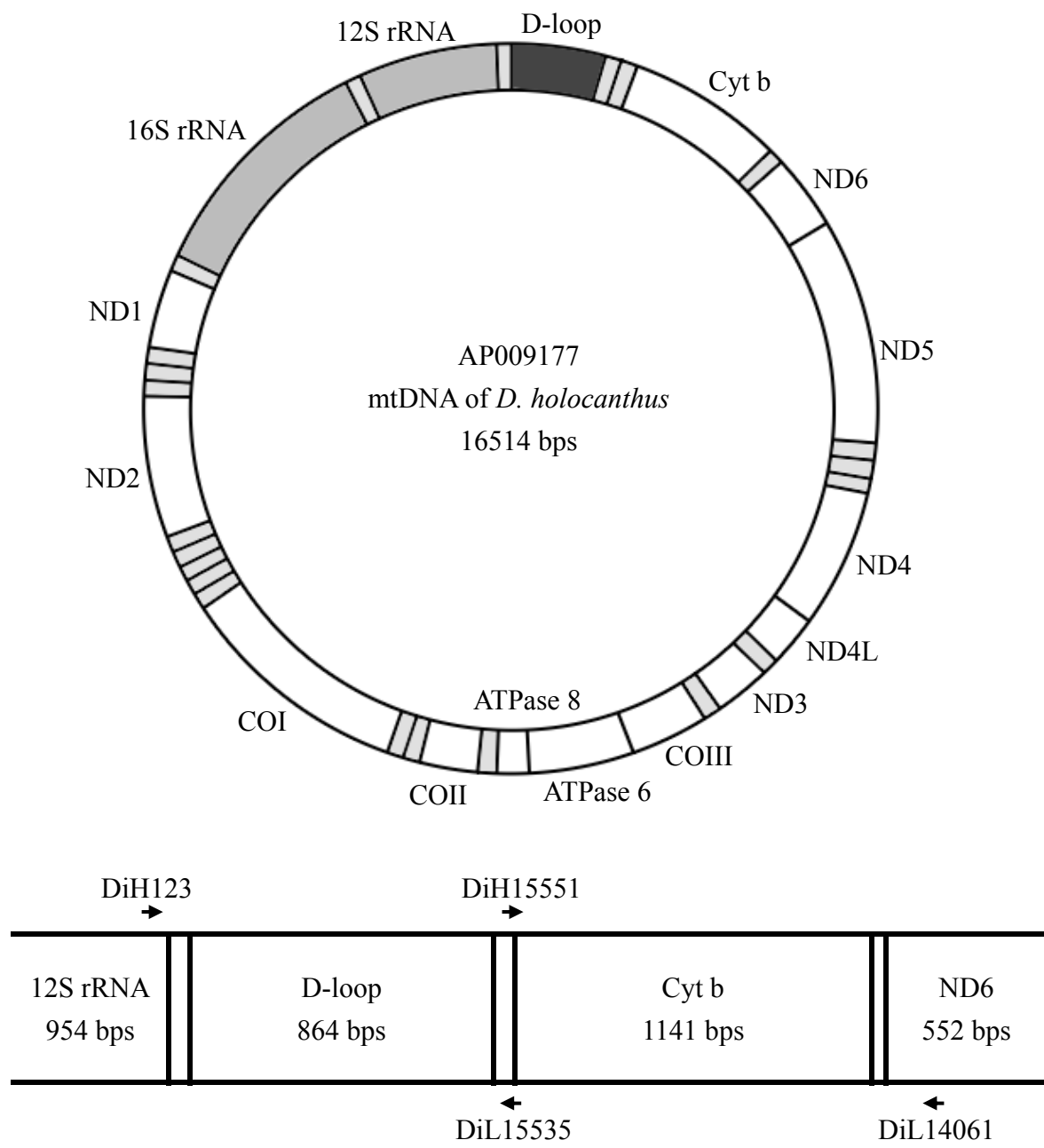
Figure 3. Above: schematic diagram of mitochondrial DNA of *D. holocanthus*, as inferred from sequence AP009177; below: schematic diagram showing position of primers sets used for PCR amplification in this study.
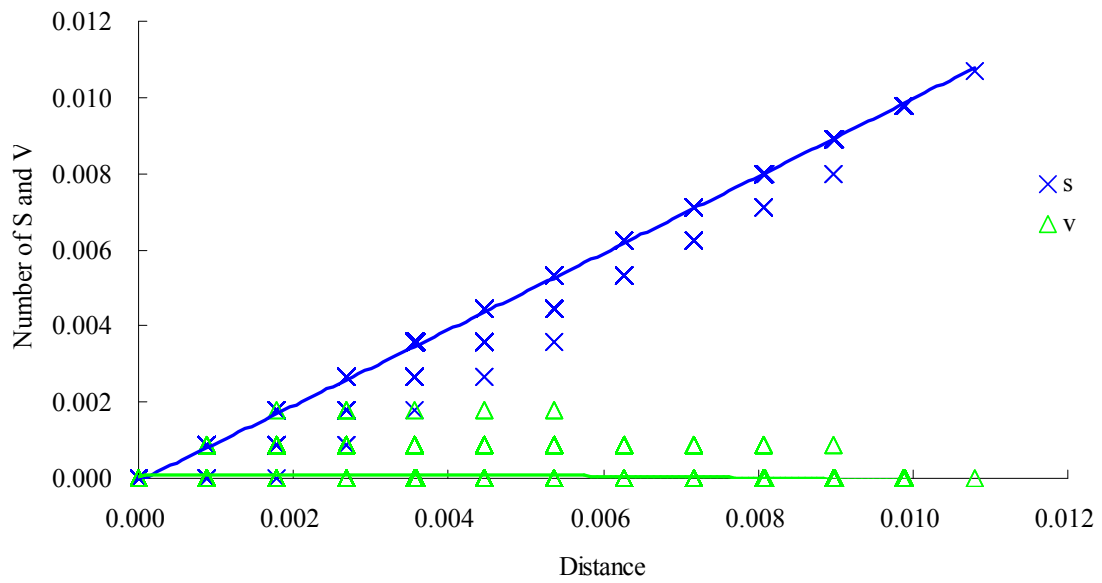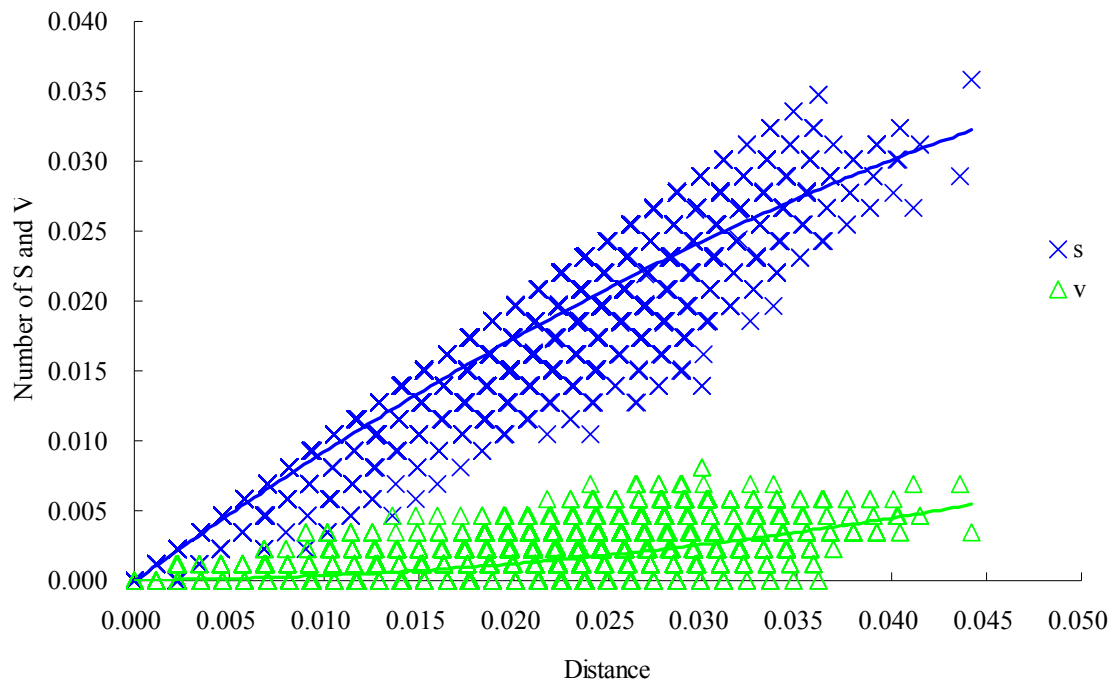
Figure 4. Plots of saturation tests, which showed estimated numbers of transition and transversion against genetic distance. Above: D-loop; below: Cyt b. S: transition; V: transversion.
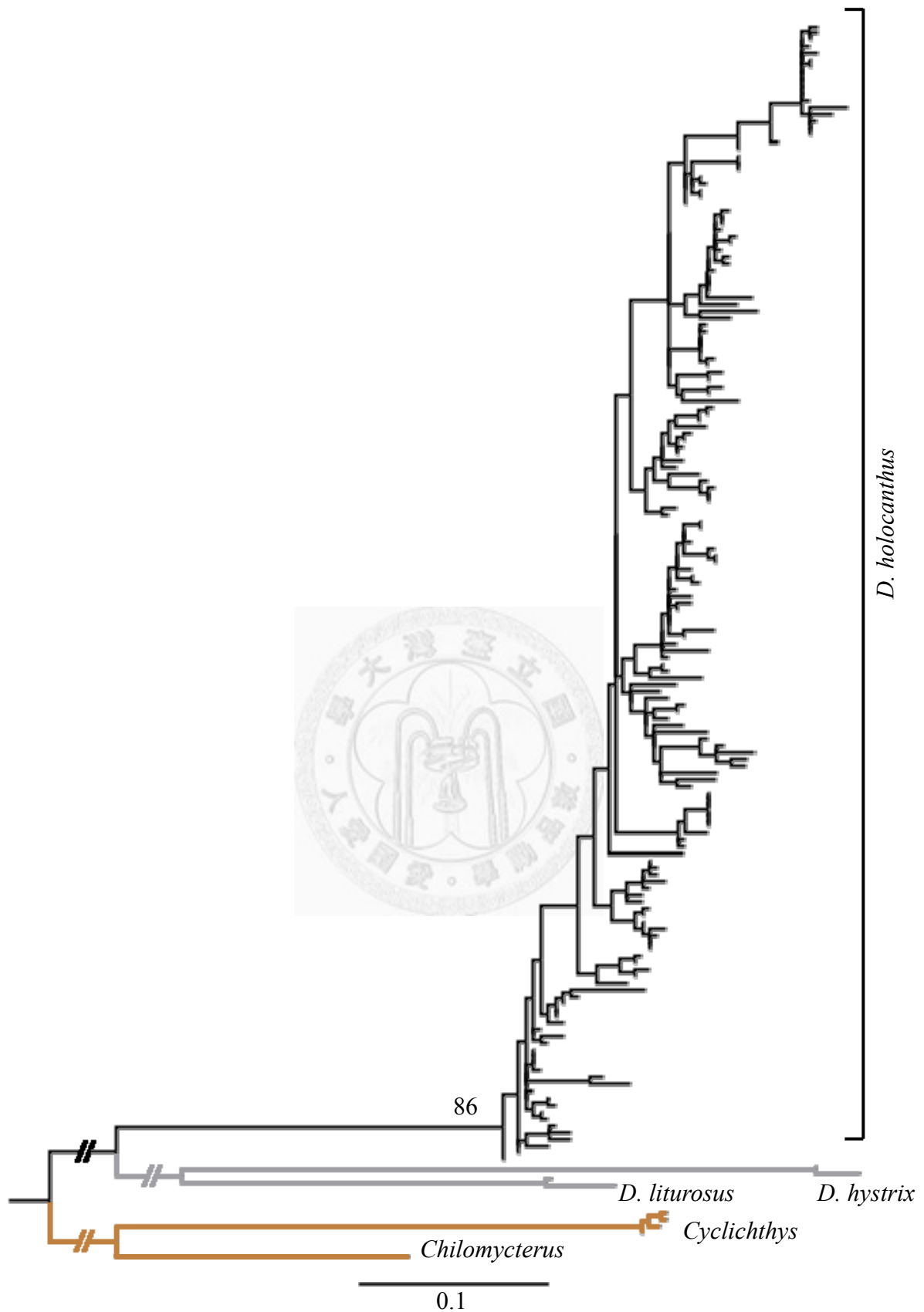
Figure 5. Phylogenetic relationship of all sampled *D. holocanthus*, reconstructed by maximum likelihood method based on D-loop sequences. Number above branch represented bootstrapping value, and values of outgroup lineages or lower than 85 were not shown.
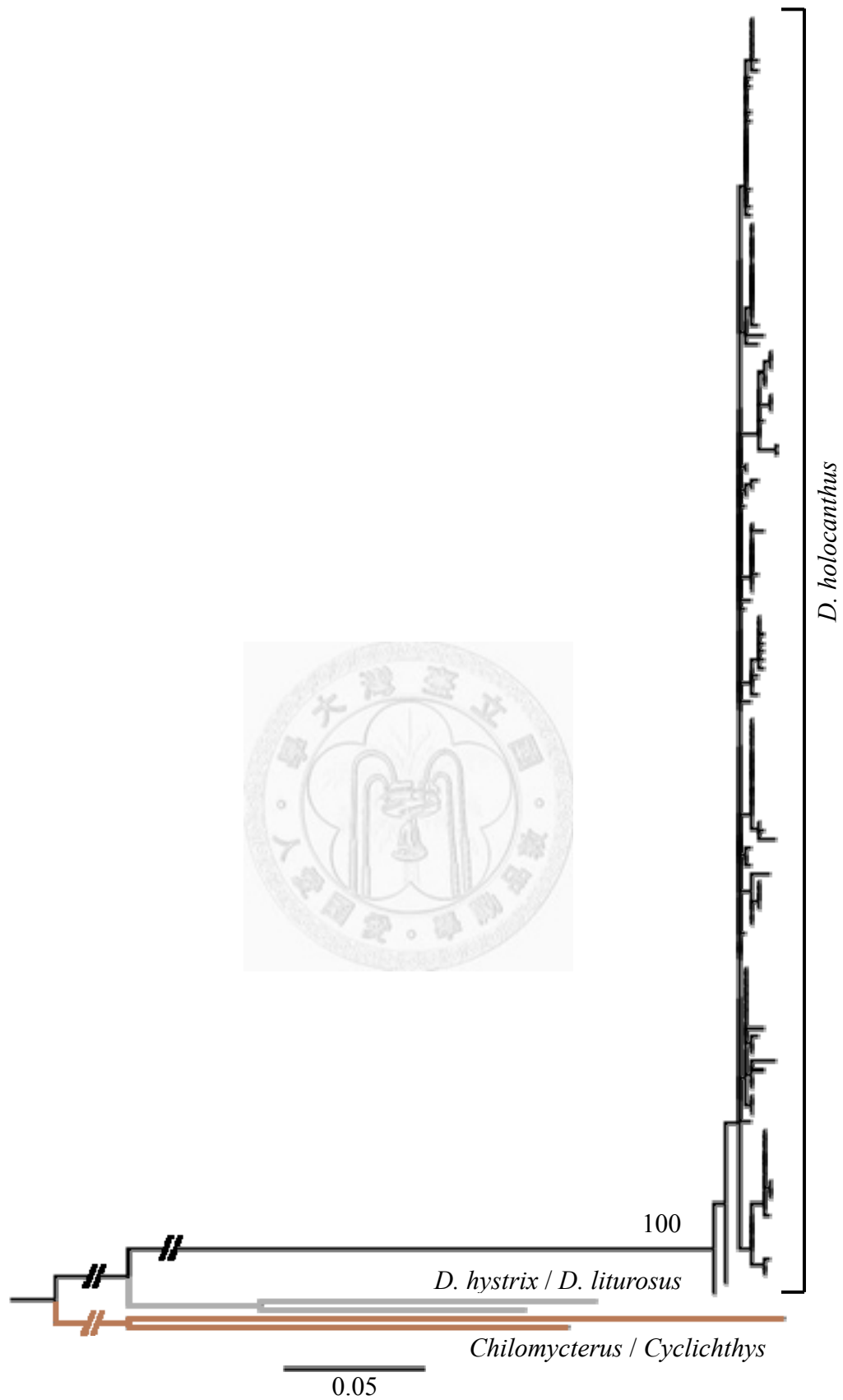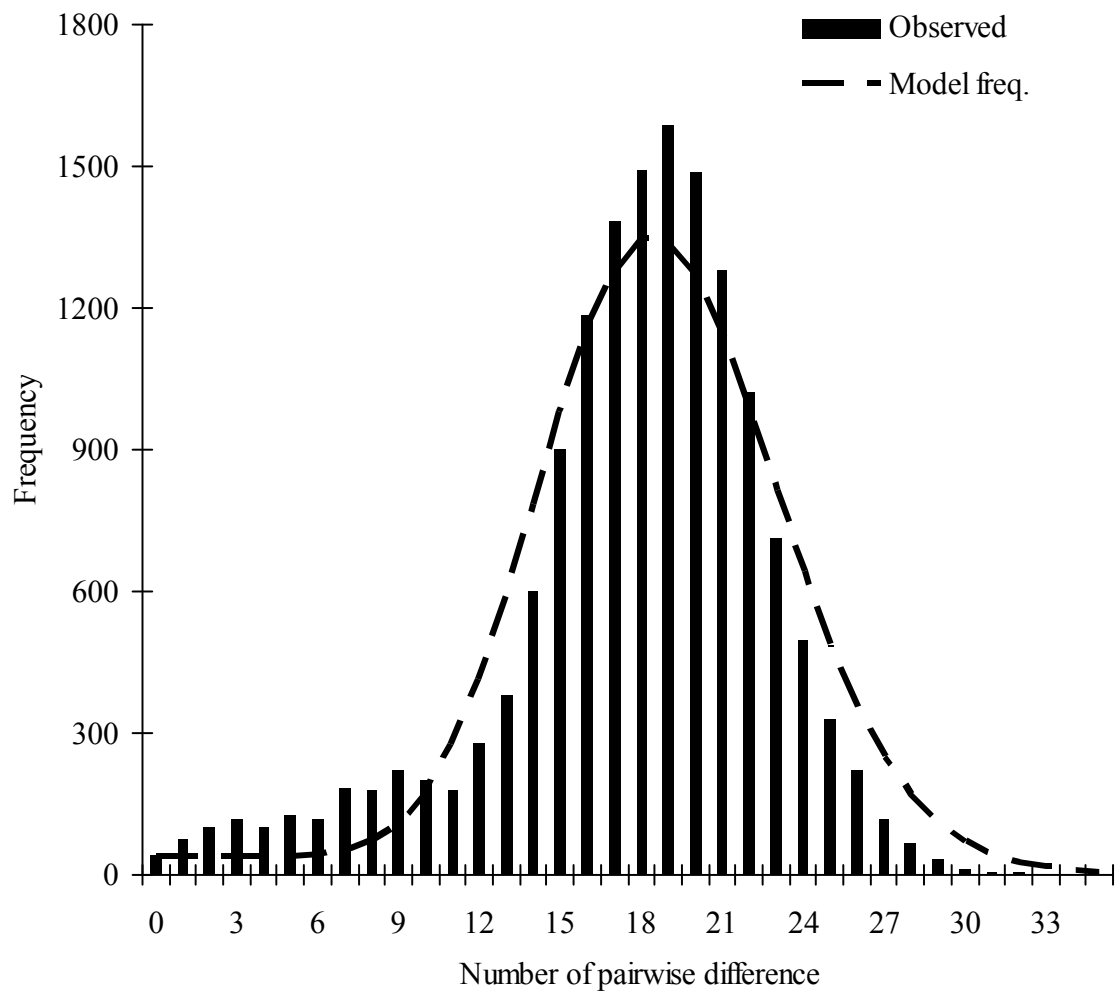
Figure 6. Phylogenetic relationship of all sampled *D. holocanthus*, reconstructed by maximum likelihood method based on Cyt b sequences. Number above branch represented bootstrapping value, and values of outgroup lineages or lower than 85 were not shown.
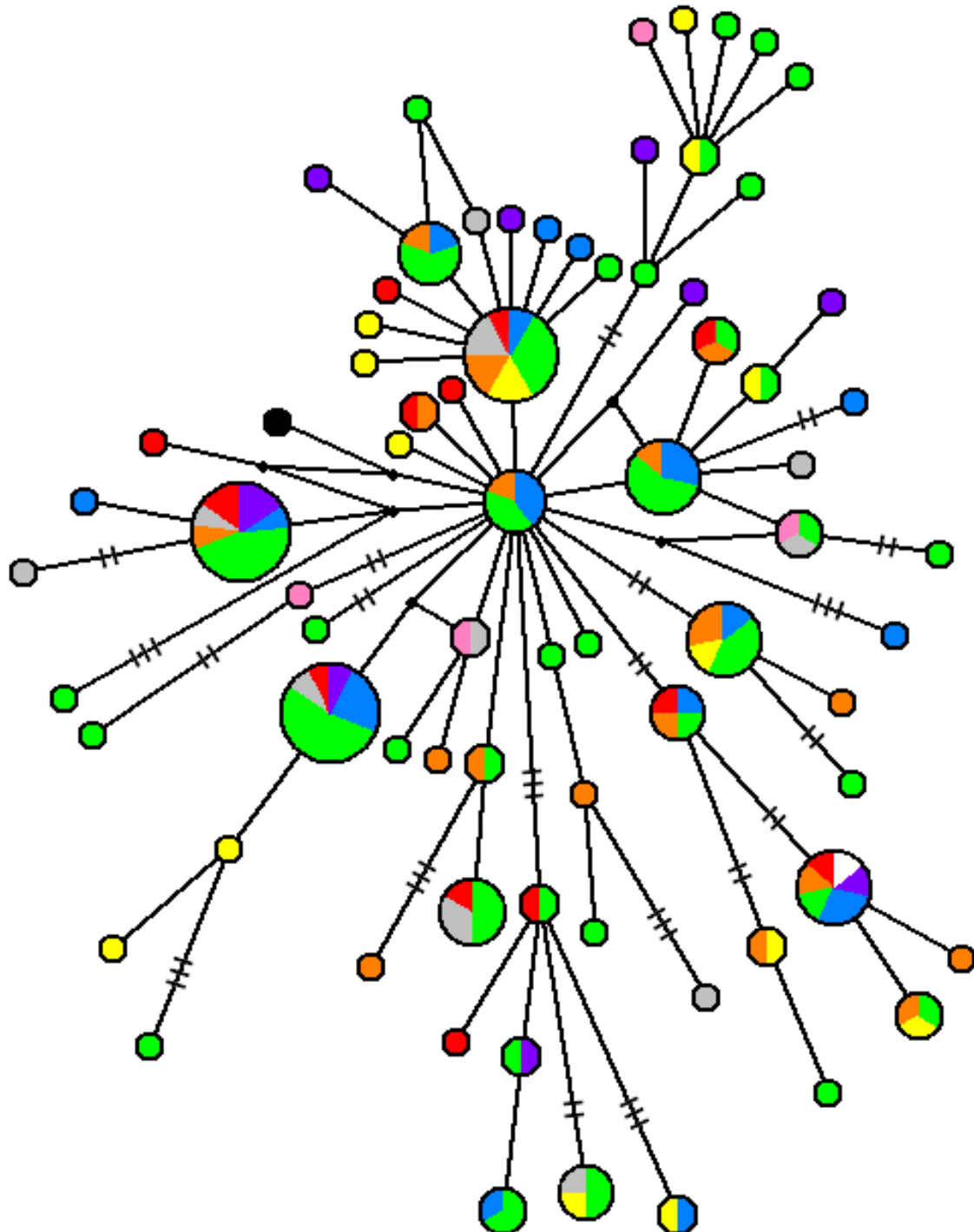
Figure 8. Mismatch distribution based on 165 D-loop sequences of sampled *Diodon holocanthus*. Black bar: observed frequencies; dash line: simulated model frequencies. SSD = 0.0020, P_{SSD} = 0.0734, r = 0.0027, P_r = 0.2201.
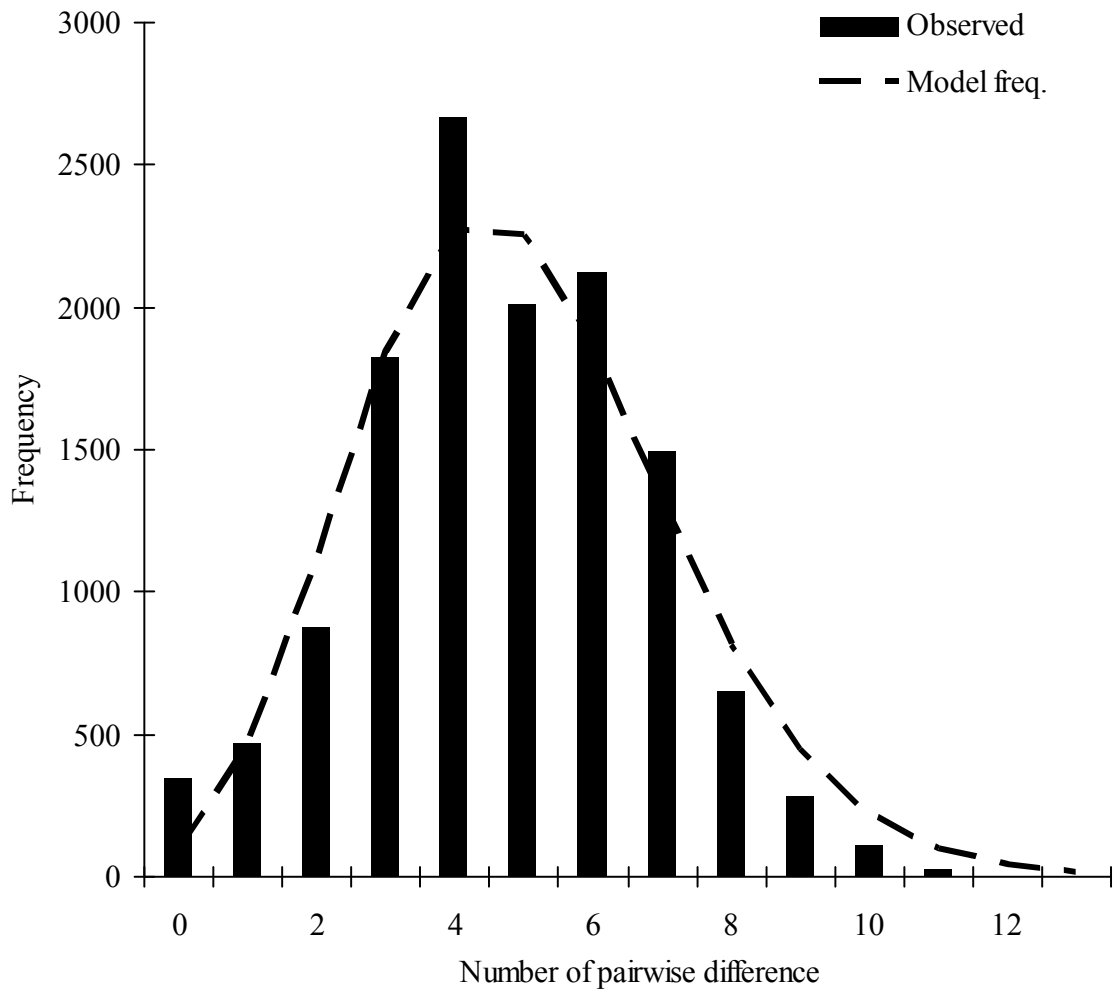
Figure 9. Mismatch distribution based on 148 Cyt b sequences of sampled *Diodon holocanthus*. Black bar: observed frequencies; dash line: simulated model frequencies. SSD = 0.0030, P_{SSD} = 0.0524, r = 0.0211*, P_r = 0.0281.
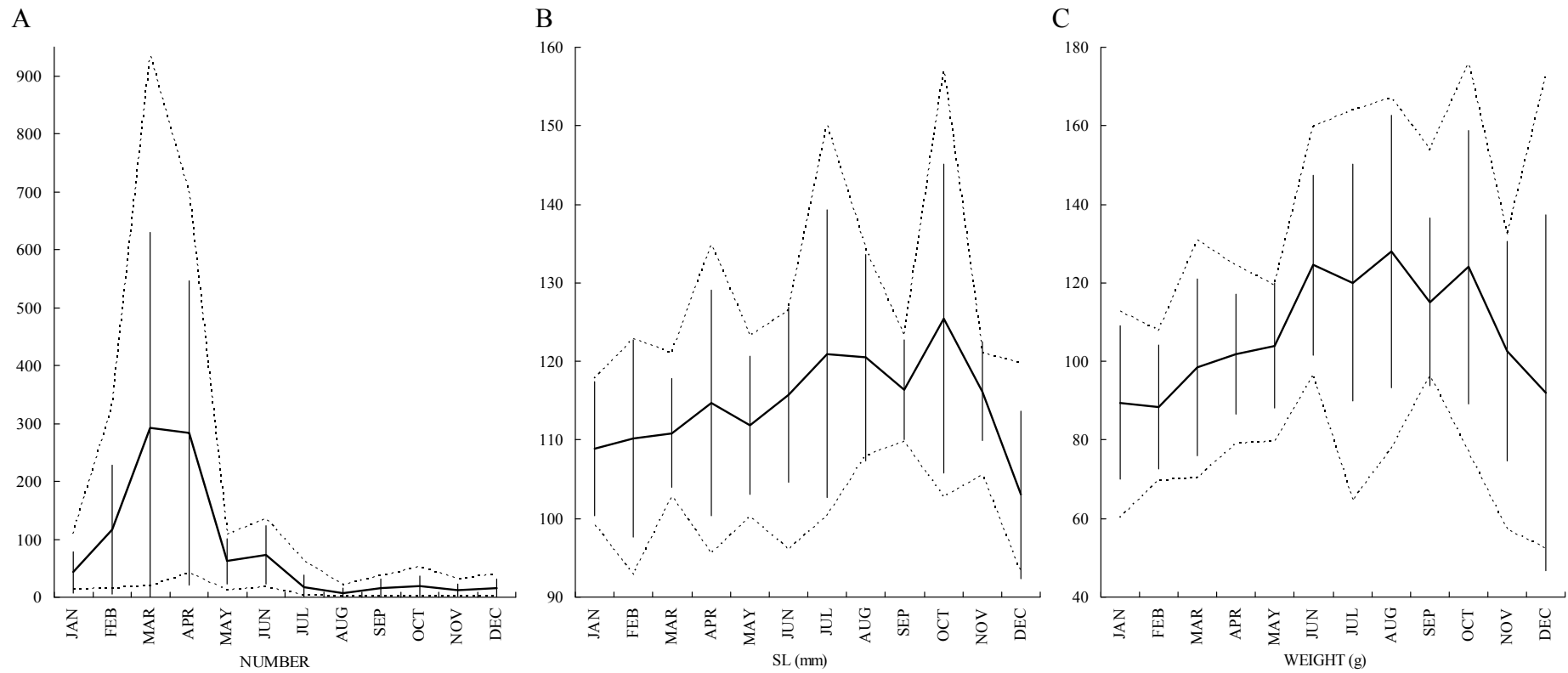
Figure 10. Monthly average number (A), length (SL, B) and weight (C) of impinged *D. holocanthus* individuals during the period from year 2000 to 2008 (Liao, unp. data). Solid line: average value; dotted line: monthly maximum or minimum values; vertical bar: standard error.
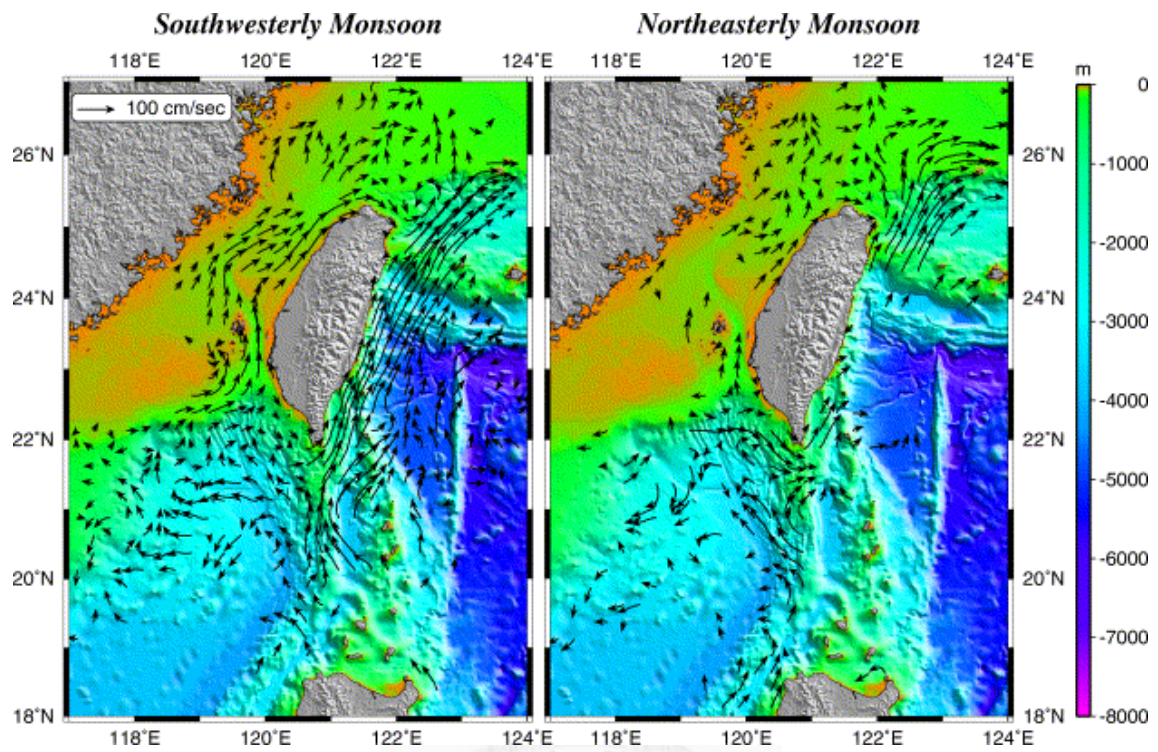
Figure 11. Surface oceanic currents in waters around Taiwan during the period when monsoon prevails, with velocity and direction. The left and right panels represent the seasons of southwesterly monsoon and northeasterly monsoon, respectively. From Liang et al. (2003).