Genetic diversity of *Dugesia* Planarians in Japan

Masato Sakai

Doctoral program in Environmental Science and Technology

Graduate School of Science and Technology

Niigata University

CONTENTS

Abstract

In the Part I, I used two sequencing methods, namely long polymerase chain reaction (PCR) and primer walking to determine the complete mitochondrial DNA (mtDNA) sequence of *Dugesia japonica* and most of the mtDNA sequence of *Dugesia ryukyuensis*. The genome of *D. japonica* contained 36 genes including 12 of the 13 protein-coding genes characteristic of metazoan mitochondrial genomes, two ribosomal RNA genes, and 22 transfer RNA genes. The genome of *D. ryukyuensis* contained 33 genes including 12 protein-coding genes, two ribosomal RNA genes, and 19 transfer RNA genes. The gene order of the mitochondrial genome from the *Dugesia* species showed no clear homology with either the Neodermata or other free-living Rhabditophora. This indicates that the platyhelminths exhibit a great variability in mitochondrial gene order. This is the first complete sequence analysis of the mitochondrial genome of a free-living Rhabditophora, which will facilitate further studies on the population genetics and genomic evolution of the Platyhelminthes.

In the Part II, I surveyed detailed genetic population structure in wild populations of *Dugesia* species in the Japanese Islands by using molecular phylogenetic analysis based on 18S rRNA type II haplotype and *coI* mitotype sequences. BI and ML tree of 18S rRNA type I and type II allocated haplotypes of *D. ryukyuensis* into one, and *D. japonica* into six genetic lineages. The divergence time based on 18S rRNA type II between *D. ryukyuensis* and *D. japonica* is estimated at 24.7-36 million years ago. NJ tree of *coI* allocated mitotypes of *D. ryukyuensis* into two (a1 and a2), and *D. japonica* into 17 genetic lineages (b-q). Since comparison of 18S rRNA type II genotypes and *coI* mitotypes among individuals demonstrated particular combinations of genotypes, I defined seven new genetic groups (R; *D.ryukyuensis*, J-I, -II, -III, -IV, -V and –IV; *D. japonica*). Furthermore,

the particular combinations of genotypes were also demonstrated in the populations where were observed two or more genotypes. This suggested that there is genetic isolation among J-I, -II, -III, -IV, -V and -VI, and existence of six sibling species in *D. japonica*.

General introduction

Geographic variation in the genetic diversity of a species reflects the influence of both historic and recent evolutionary processes.

The flatworms, known in scientific literature as Platyhelminthes, are a phylum of relatively simple bilaterian, unsegmented, soft-bodied invertebrate animals. They comprise over 22,000 described species. Unlike other bilaterians, they have no body cavity, and no specialized circulatory and respiratory organs, which restrict them to flattened shapes that allow oxygen and nutrients to pass through their bodies by diffusion.

Traditional classifications subdivide the Platyhelminthes into Turbellaria, Monogenea, Trematoda, and Cestoda. Turbellarians are free-living flatworms, which, unlike other classes of Platyhelminthes, are not parasitic. Parasitic flatworms (Monogenea, Trematoda, and Cestoda), named as the "Neodermata" (Ehlers 1985, 1986), share common features including shedding of the epidermis at the end of the larval phase, post-larval stages with a syncytial neodermis, protonephridium with a two-cell weir, and spermatozoa with completely incorporated ciliary axonemata. Molecular phylogenetic studies indicate that Neodermata are monophyletic (Baguñà and Riutort 2004a). In contrast, Turbellaria, which mainly constitute the free-living flatworms, are considered to be an invalid class, because detailed morphological analyses of their anatomical features suggest that they are not monophyletic (Ehlers 1985, 1986). The common denominator of all phylogenetic schemes recently proposed for the Platyhelminthes is the recognition of three clearly monophyletic groups: the Acoela and Nemertodermatida ("Acoelomorpha," Ehlers 1985, 1986); the Catenulida; and all other turbellarian orders, together with the parasitic classes ("Rhabditophora," Ehlers 1985, 1986). Molecular phylogenetic studies have shown that Turbellaria are polyphyletic. The Catenulida are the sister group of Rhabditophora within the Lophotrochozoa (Larsson and Jondelius 2008),

whereas the Acoelomorpha are a sister group of the bilaterians (Baguñà and Riutort 2004b; Jondelius et al. 2002; Ruiz-Trillo et al. 1999, 2004), or alternatively, both the acoelomorphs and *Xenoturbella* are within the Deuterostomes (Philippe et al. 2011). At present, the redefined Platyhelminthes consist of two monophyletic sub-groups, the Catenulida and the Rhabditophora. Furthermore, little is known about the intraspecific and interspecific phylogenetic relationships within free-living the Rhabditophora (Marta et al., 2008, Lázaro et al., 2009).

Freshwater planarians of the genus *Dugesia* (Order Tricladida), which are a free-living the Rhabditophora, comprise over 70 described species with a wide distribution, viz. the Afrotropical, Palearctic, Oriental, and Australian biogeographic regions. Of these 70 species, two species are mainly distributed in the Japanese Islands. *Dugesia japonica* is habitats widely in the Far East (the Japanese Islands, Taiwan, the Korean Peninsula, China, and Primorskiy in Russia) (Kawakatsu et al., 1995). *Dugesia ryukyuensis* is recorded in the Southwest Islands of Japan (Nansei Shoto) and the lowland areas in the Kyushu distinct on the East China Sea (Tamura et al., 1998).

In the Part I, I used two sequencing methods, namely long polymerase chain reaction (PCR) and primer walking to determine the complete mitochondrial DNA (mtDNA) sequence of *Dugesia japonica* and most of the mtDNA sequence of *Dugesia ryukyuensis*. The gene order of the mitochondrial genome from the *Dugesia* species showed no clear homology with either the Neodermata or other free-living Rhabditophora. This indicates that the platyhelminths exhibit a great variability in mitochondrial gene order.

In the Part II, I surveyed detailed genetic population structure in wild populations of *Dugesia* planarians in the Japanese Islands by using molecular phylogenetic analysis based on 18S rRNA type II haplotype and *coI* mitotype sequences. A large number of 18S rRNA type II haplotype and mitotypes were identified. The distribution patterns of each of 18S rRNA

type II genotype and *coI* mitotype demonstrated geographical associations. Since comparison of 18S rRNA type II genotypes and *coI* mitotypes among individuals demonstrated particular combinations of genotypes, seven new genetic groups have been defined (R; *D.ryukyuensis*, J-I, -II, -III, -IV, -V and –IV; *D. japonica*).

Part I. The complete mitochondrial genome of *Dugesia japonica* (Platyhelminthes; Order Tricladida)

Introduction

Parasitic flatworms (Monogenea, Trematoda, and Cestoda), named as the "Neodermata" (Ehlers 1985, 1986), share common features including shedding of the epidermis at the end of the larval phase, post-larval stages with a syncytial neodermis, protonephridium with a two-cell weir, and spermatozoa with completely incorporated ciliary axonemata. Molecular phylogenetic studies indicate that Neodermata are monophyletic (Baguñà and Riutort 2004a). In contrast, Turbellaria, which mainly constitute the free-living flatworms, are considered to be an invalid class, because detailed morphological analyses of their anatomical features suggest that they are not monophyletic (Ehlers 1985, 1986). The common denominator of all phylogenetic schemes recently proposed for the Platyhelminthes is the recognition of three clearly monophyletic groups: the Acoela and Nemertodermatida ("Acoelomorpha," Ehlers 1985, 1986); the Catenulida; and all other turbellarian orders, together with the parasitic classes ("Rhabditophora," Ehlers 1985, 1986). Molecular phylogenetic studies have shown that Turbellaria are polyphyletic. The Catenulida are the sister group of Rhabditophora within the Lophotrochozoa (Larsson and Jondelius 2008), whereas the Acoelomorpha are a sister group of the bilaterians (Baguñà and Riutort 2004b; Jondelius et al. 2002; Ruiz-Trillo et al. 1999, 2004), or alternatively, both the acoelomorphs and *Xenoturbella* are within the Deuterostomes (Philippe et al. 2011). These studies have also concluded that the redefined Platyhelminthes consist of two monophyletic sub-groups, the Catenulida and the Rhabditophora.

The animal mitochondrial genome provides a large set of orthologous sequence data that are often used in phylogenetic analyses from the population to the phylum level. Comparisons of complete mitochondrial genome sequences, including comparisons of gene order, are powerful tools for addressing phylogenetic relationships (Boore and Brown 1998).

Complete mitochondrial DNA (mtDNA) sequences are available for 30 parasitic species of Platyhelminthes (Perkins et al. 2010). However, only a partial mitochondrial genome (6.7 kb) of the free-living Rhabditophora (order Microstomidae) *Microstomum lineare* has been reported (Ruiz-Trillo et al. 2004), and no complete sequence data are available for any of the free-living Rhabditophora and Catenulids.

I determined the complete mtDNA sequence of the free-living Rhabditophora, *D. japonica* (order Tricladida), which is a common fresh-water planarian, and an almost-complete mtDNA sequence of *D. ryukyuensis* as supporting data. Molecular phylogenetic analysis based on 18S rDNA sequences alone (Littlewood et al. 2001), or on the combined 18S and 28S rDNA sequences (Lockyer et al. 2003), has suggested that the Tricladida together with other groups such as Prolecitophora and Rhabdocoela, among others, are a sister group of the Neodermata. However, the gene order of the mitochondrial genome of the *Dugesia* species showed no clear homology with either the Neodermata or *M. lineare* mitochondrial gene orders. This suggests that the platyhelminths show a great variability in mitochondrial gene order.

Materials and Methods

Animals

The freshwater planarians *Dugesia japonica* were collected from Gosen, Niigata and Unnan, Shimane, Japan and *Dugesia ryukyuensis* from Naha, Okinawa, and Shibushi, Kagoshima, Japan, respectively. The phenol-chloroform method was used to isolate genomic DNA from individual organisms after they had been housed without food for a week. Individual organisms were then crushed in 200 ml lysis buffer (0.1 M EDTA [pH 8.0], 0.05 M Tris-HCl [pH 8.0], 0.1 M NaCl, and 1% SDS) with 40 mg Proteinase K (TaKaRa, Shiga, Japan) using a pestle fitting a 1.5-ml Eppendorf tube, and incubated at 55°C for 2 h. An equal volume of phenol/chloroform was added to the tubes. After centrifugation at 15,000 rpm for 5 minutes and transferring the aqueous phase to a new tube, an equal volume of isopropanol acetate was added. After a second centrifugation step, the resultant DNA pellet was washed with 70% ethanol, dried, and suspended in 150 ml Tris-EDTA (TE) buffer (1 mM EDTA, 10 mM Tris [pH 8.0]).

Long polymerase chain reaction (PCR) and sequencing by primer walking

The partial *coI* region of *D. japonica* from Gosen was determined using previously published primers (Bessho et al. 1992a, b), and was confirmed to be a single sequence. Three long-PCR primers were designed (Table 1). The first long PCR was performed by using the primer set LA_f01 and LA_rev01. PCR was performed in a 50 µl reaction mixture containing 30 µl water, 10 µl 5 × PrimeSTAR GXL buffer, 1.0 µl PrimeSTAR GXL DNA polymerase (5 U/µl) (TaKaRa, Shiga, Japan), 4 mM dNTPs, 4.0 µM of each primer, and 1.0 µl of template DNA (50 ng). The thermal cycle profile was set as follows: 35 cycles of 98°C for 10 s and 68°C for 10 min. A 1.0 µl aliquot of the PCR

product was used for the second long PCR (nested PCR), which was conducted under the same conditions and thermal cycles, except for the use of LA_f02 primers instead of LA_f01. The PCR product was analyzed by electrophoresis on an agarose gel and then purified by the phenol/chloroform method. The purified product was then sequenced using the primer-walking protocol. Long-PCR primers were used for the initial sequencing of the 5' and 3' ends of the PCR products. The remaining nucleotide sequences were determined using 44 newly designed sequencing primers (Table 1). The obtained sequences around 800 bp in length, and each sequence overlapped the next contig by over 100 bp. Sequencing was performed using the BigDye terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) in a 3130 DNA Analyzer (Applied Biosystems). Sequences were assembled using ClustalW (Thompson et al., 1994) implemented in MEGA4 software package (Tamura et al. 2007), and the circular map was constructed using the Vector NTI v10.1.1 software package (Invitrogen, Carlsbad, CA, USA).

The partial mitochondrial genomes of *D. japonica* from Unnan and *D. ryukyuensis* from Naha and Shibushi were sequenced after the partial *coI* region had been determined and confirmed to be a single sequence. PCR performed in a 10-μl reaction mixture containing 6.7 μl water, 1.0 μl 5 × Takara EX *Taq* buffer, 0.025 μl EX *Taq* polymerase, 4 mM dNTP, 2.5 μM of each primer, and 1.0 μl of template. The thermal cycle profile was 35 cycles of 95°C for 30 s, 60°C for 30 s and 72°C for 30 s. Sequencing was performed using the BigDye terminator v.3.1 Cycle Sequencing Kit. Long PCR was performed using the primer sets cytb_f01 and LA_rev07, LA_f07 and rrnL_rev01, and rrnL_f02 and cytb_rev02 (Table 1). PCR were performed in a 50 μl reaction mixture containing 30 μl water, 10 μl 5 × PrimeSTAR GXL buffer, 1.0 μl PrimeSTAR GXL DNA polymerase (5 U/μl), 4 mM dNTPs, 4.0 μM of each primer, and 1.0 μl (50 ng) of template. The thermal cycle profile was set as follows: 35 cycles of 98°C for 10 s and

68°C for 10 min. The PCR products were analyzed by electrophoresis on an agarose gel and then purified by the phenol/chloroform method. Each of the PCR products was sequenced by primer walking. Long-PCR primers were used for the initial sequencing of the 5' and 3' ends of the PCR products. The remaining nucleotide sequences were determined using 32 newly designed sequencing primers (Table 1). The sequences reported in this paper have been deposited in the DNA Data Bank of Japan (DDBJ) under accession numbers AB618487 and AB618488.

Annotation of sequences

The sequence identity of the open reading frames was verified by using BLAST programs in the National Center for Biotechnology Information (NCBI) database, and individual genes were aligned with published mitochondrial genomes on the basis of the start codons (ATN, GTG, TTG, and GTT) and stop codons (TAG or TAA). Protein-coding regions were translated with the help of echinoderm and flatworm mitochondrial codes. Large (*rrnL*) and small (*rrnS*) ribosomal RNA (rRNA) genes were initially identified by BLAST search. The secondary structure of *D. japonica* was determined mainly by eye. The *rrnL* and *rrnS* genes of *D. ryukyuensis* were identified by comparing their homology with that of *D. japonica*. The transfer RNA (tRNA) genes were identified manually by comparison with flatworm anticodon sequences and their putative tRNA-like secondary structures. AT and GC skew were calculated according to the formulae: AT skew = $(fA - fT)/ (fA + fT)$ and GC skew = $(fG - fC)/ (fG + fC)$.

Prediction of stem-loop secondary structure

Stem-loop secondary structures in the non-coding region were predicted by using the program CentroidFold (http://www.ncrna.org/centroidfold) by selecting the default option.

Nucleotide sequence divergences of protein coding and ribosomal RNA genes

Intraspecific and interspecific nucleotide divergences of protein coding and ribosomal RNA genes of *D. japonica* and *D. ryukyuensis* were calculated using the Tamura-Nei model with gamma distribution.

Results

General features of the mitochondrial genomes of *D. japonica* and *D. ryukyuensis*

The circular complete mitochondrial genome of *D. japonica* from Gosen consisted of 17,799 bp, whereas the nearly complete mitochondrion genome of *D. ryukyuensis* from Naha and Shibushi had 17,105 bp (Fig. 1, Table 2) and 15, 596 bp, and that of *D. japonica* from Unnan 14, 154 bp.

The genome of *D. japonica* from Gosen contained 36 genes that included 12 of the 13 protein-coding genes characteristic of metazoan mitochondrial genomes, 2 ribosomal RNA (rRNA) genes, and 22 tRNA genes. The genome of *D. ryukyuensis* from Naha contained 33 genes that included 12 protein-coding genes, 2 ribosomal RNA genes, and 19 transfer RNA genes. The gene arrangements were common to both the species (Fig. 2). All the genes were transcribed from the same strand, with the exception of *trnC* in *D. japonica* from Gosen.

The genome of *D. japonica* from Gosen had two large non-coding regions, located between *trnA* and *trnC* and between *trnC* and *cytb*. *D. ryukyuensis* from Naha had two large non-coding sequences located between *trnT* and *cytb*. However, we could not determine the correct size of either sequence, because they were interrupted by a repeat region.

The genes contained in the mitochondrial genomes of *D. japonica* from Gosen and *D. ryukyuensis* from Naha are listed in Table 2. The compositions of the respective mitochondrial genomes were 22.7% and 21.7% A, 53.0% and 51.6% T, 15.6% and 16.0% G, and 9.1% and 10.7% C, with a total A + T content of 75.7% and 73.3%. Both species were very rich in A and T. The AT skew and GC skew (Perna and Kocher 1995) of all the plus-strand sequences were –0.40 and 0.26 for *D. japonica* and –0.41 and 0.20 for *D. ryukyuensis*. (Skew values range from +1 to –1; the value is 0 if the strands have no skew).

Protein-coding genes

BLAST searches identified all protein-coding genes except *atp8* in both *Dugesia* species. *atp8* is absent in the mitochondrial genomes of the Platyhelminthes, the Chaetognatha, and most nematodes (Gissi et al. 2008). Only *nd4l* overlapped with other protein-coding genes in *D. japonica* from Gosen (40 bp with *cytb* and 32 bp with *na4*) and *D. ryukyuensis* (43 bp with *cytb* and 32 bp with *na4*).

The *coII* sequence of *D. ryukyuensis* from Naha did not have a terminal codon (TAG or TAA); instead, it had sequences including a short tandem repeat unit (5′ GAT GAT GTA ACA CCY CCT ATT GAT GAA GAT GTG GTT TTA CCT CCT ATT GTT ATT). The *coII* terminal of *D. ryukyuensis* was identified through its homology with the *coII* terminal of *D. japonica*.

Four and five start codons were used in the *D. japonica* from Gosen and *D. ryukyuensis* from Naha protein-coding genes, respectively. In *D. japonica* from Gosen, ATG was found in *coII*, *nd4*, *nd4l,* and *nd5*; ATT in *cytb*, *nd1*, and *nd5*; ATA in *nd6*; and TTG in *coI*, *coIII*, *atp6*, and *nd3*. In *D. ryukyuensis* from Naha, ATG was found in *coII*; ATT in *nd1* and *nd6*; ATA in *nd2*; TTG in *cytb*, *coI*, *coIII*, *atp6*, *nd3,* and *nd4l*; and GTG in *nd4* and *nd5*. All protein-coding genes terminated with the codon TAG or TAA, with the exception of the *coII* gene of *D. ryukyuensis* from Naha that ended with T, as found in many metazoan mitochondrial genomes.

The utilization of an unusual genetic codon (UAA = Tyr) was reported in *D. japonica* (Bessho et al. 1992 a, b); however, this finding was not observed in the *D. japonica* and *D. ryukyuensis* protein-coding genes in the present study.

tRNA genes

Twenty-two tRNA genes were identified in the *D. japonica* from Gosen mtDNA sequence, and their putative secondary structures were determined

(Fig. 2). Because the initial software search failed to identify any tRNA genes, the tRNA sequence was determined by searching for the YUXXXRN structure of the anti-codon loop by eye from the entire sequence and/or sequences that were homologous between *D. japonica* and *D. ryukyuensis*, using the Clustal W program. Most tRNA genes of *D. japonica* showed mismatches of one to several nucleotides, with shortened stems or enlarged loops. The DHU stem was missing in *trnQ* and *-T*. The TψC-stem was missing in *trnV*. Two tRNA genes overlapped with other protein-coding genes, namely *trnV* (14 bp with *nd1*) and *trnP* (15 bp with *coII*).

Nineteen tRNA genes of *D. ryukyuensis* from Naha were identified; however, *trnA, -C*, and *-E* were not found. The DHU stem was missing in *trnI, -K, -S1, -S2*, and *-T* (data not shown).The TψC-stem was missing in *trnD* (data not shown). Three tRNA genes overlapped with other protein-coding genes, namely *trnV* (3 bp with *atp6* and 20 bp with *na1*) and *trnL* (16 bp with *nd2*).

rRNA genes

Two rRNA genes of *D. japonica* were identified, and their putative secondary structures were determined (Fig. 3 A, B). Although BLAST searches were performed, the sequence identity of the rRNA genes with those of other organisms was very low. The secondary structures of the *D. japonica* rRNA genes were then confirmed, mainly by eye, in accordance with the protocol described by Gutell et al. (2000a, b). Two rRNA genes of *D. ryukyuensis* from Naha were identified by their high level of sequence homology with those of *D. japonica* (data not shown).

Non-coding regions

D. japonica from Gosen had two large non-coding regions, NC1-J and NC2-J, which were 1390 bp and 1927 bp long, respectively. NC1-J was found to be a tandem repeat region consisting of a 208-bp unit. NC2-J was

AT rich, consisting mainly of AT or ATT repeats, with a total A + T content of 88.7%, and was able to fold into a stem-loop secondary structure (Fig. 4). A total of 24 short non-coding regions were identified ranging from 1 bp to 330 bp in length. The non-coding regions of *D. japonica* from Gosen did not show any sequence homology with mtDNA non-coding regions of other flatworms.

D. ryukyuensis from Naha also had large non-coding regions, NC1-R and NC2-R, consisting of approximately 849 bp and 212 bp, respectively. NC1-R and NC2-R could not be sequenced correctly, because they had a large number of tandem repeats. NC1-R was found to be a tandem repeat region consisting of a 166-bp unit, with no homology to the *D. japonica* unit. NC2-R was AT rich, consisting mainly of AT or ATT repeats. A total of 20 short non-coding regions were identified, ranging from 1 bp to 484 bp in length.

Gene order

Although the gene order in *D. japonica* and *D. ryukyuensis* was similar, the points on the *D. ryukyuensis* genetic map in which *trnA*, *-C*, and *-E* were not found were different, and the positions of NC1 and NC2 in *D. japonica* and *D. ryukyuensis* were reversed (Fig. 1). I compared the sequences of *D. japonica* and *D. ryukyuensis* to the complete sequences of *Trichobilharzia regenti* (Trematoda; NC_009680), *Gyrodactylus salaris* (Monogenea; NC_008815), and *Hymenolepis diminuta* (Cestoda; NC_002767), and to the partial sequences of *M. lineare* (Microstomidae; AY228756) (Fig. 5). The gene order of the Neodermata appears relatively well conserved (Park et al. 2007). However, there was no obvious sequence homology between the *Dugesia* species and the Neodermata. The order of the protein-coding and rRNA genes revealed three gene blocks at *cytb-nd4l-nd4*, *nd6-nd5-coIII*, and *rrnS-rrnL*, which were conserved in *Dugesia* and in the Neodermata. There was almost no sequence homology between *Dugesia* and *M. lineare*,

except for the conservation of the *rrnS-rrnL* gene block. The *Dugesia* species showed no clear homology with any other mitochondrial gene order published to date.

Intraspecific and interspecific nucleotide sequence divergences of protein coding and ribosomal RNA genes of *Dugesia speices*

The intraspecific and interspecific nucleotide divergences within/between *D. japonica* and *D. ryukyuensis* were shown in Table 3. The overall intraspecific nucleotide divergence of *D. japonica* was 0.42, and *D. ryukyuensis* 0.08. The overall interspecific nucleotide divergence between *D. japonica* and *D. ryukyuensis* was 0.54.

Discussion

Using long-PCR and primer-walking sequencing methods, the complete mtDNA sequence of *D. japonica* and most of the mtDNA sequence of *D. ryukyuensis* were determined.

A total of 12 protein-coding genes were identified in *D. japonica* and *D. ryukyuensis*; however, utilization of the unusual genetic codon (UAA = Tyr), which was previously considered to be rarely used, was not observed in this study. Further analysis of individual mtDNAs obtained from *D. japonica* and *D. ryukyuensis* is necessary to reveal the presence or absence of this unusual genetic codon.

Twenty-two tRNA genes were identified in the *D. japonica* mtDNA sequence, whereas no sequences resembling the genes for *trnA, -C*, and *-E* could be detected in the other genes of *D. ryukyuensis*. In *D. ryukyuensis*, two YUXXXRN structure sequences containing the *trnE* anticodon were found in the region between *nd2* and *trnN* (12915–12921 and 12831–12837; reverse direction), which was the same region where putative *trnE* was found, but the flanking sequences of this region did not form the cloverleaf secondary structure. Instead, a candidate anticodon sequence of *trnE* was found in *coI* (2950–2956), and it was capable of forming the cloverleaf secondary structure. Because the candidate positions for *trnA* and *-C* were not found in *D. ryukyuensis*, it is possible that the yet unidentified region between NC1 and NC2 might contain these genes. These results suggest that the gene order has diverged in the genus *Dugesia*.

D. japonica and *D. ryukyuensis* had two large non-coding regions. The NC2 regions were considered to be homologous in these species, because they were consistently AT-rich and consisted mainly of AT or ATT repeats. However, the NC1 regions were not considered homologous in *D. japonica* and *D. ryukyuensis*. It is reported that typical control regions are not readily identifiable within the mtDNA of flatworms (Huyse et al. 2007). Huyse et al.

(2007) further suggested that a smaller non-coding region might represent the control region, which have the typical feature of a control region (Wolstenholme 1992). The NC2-J sequence showed no homology with the non-coding mtDNA region of other flatworms; however, it exhibited a higher AT content (88.7%) and contained tandem repeats and a stem-loop secondary structure. Hence, NC2 may be considered as the control region in *D. japonica* and *D. ryukyuensis*. The *D. japonica* repeat unit in NC1 showed no homology to the *D. ryukyuensis* unit, and the positions of NC1 and NC2 in *D. japonica* and *D. ryukyuensis* were reversed. This suggests that the NC1 regions occurred independently.

The mitochondrial gene order of *Dugesia* is markedly different from that of the Neodermata and *M. lineare*. Only the relative position of *rrnS* and *rrnL* is conserved between *Dugesia*, the Neodermata, and *M. lineare*, albeit with changes in tRNA genes as follows: *trnY*, *-G*, and *-S1* (*Dugesia*) instead of *trnC* (Neodermata) or *-S2* (*M. lineare*). Furthermore, the relative positions of *nd6*, *nd5*, and *coIII* were conserved in *Dugesia* and the Neodermata, albeit with changes in tRNA genes between *nd6* and *nd5* as follows: no tRNA (*Dugesia*) instead of *trnY*, *-L1*, *-S2*, *-L2*, and *-R* (Neodermata), and between *nd5* and *coIII* as follows: *trnS2*, *-D*, and *-R* (*Dugesia*) instead of *trnG* (Neodermata). These results indicate that there is great variability in the mitochondrial genome structure within the Rhabditophora due to the complex history of gene rearrangements.

While genetic diversity was low level within *D. ryukyuensis*, it was high within *D. japonica*. The overall genetic distance based on the protein-coding and ribosomal RNA genes between *G. salaris* and *G. derjavinoides* (Monogenea) amounted to 22.5% (Huyse et al., 2008). The mean genetic distance of protein-coding genes between *S. haematobium* and *S. spindale* (Cestoda) was 29% (Zarowiecki et al., 2008). In comparison with these values, intraspecific variation within *D. japonica* is very high. Therefore, it is considered that at least two subspecies exist in *D. japonica*.

This is the first complete sequence analysis of the mitochondrial genome of a free-living Rhabditophora. As more mitochondrial genomes are characterized for a wider taxonomic coverage of the phylum Platyhelminthes, the phylogenetic and taxonomic utility of these features can be evaluated.

Part II. Geographic diversity of *Dugesia* Planarians in Japan

Introduction

Freshwater planarians of the genus *Dugesia* comprise over 70 described species with a wide distribution, viz. the Afrotropical, Palearctic, Oriental, and Australian biogeographic regions. Of these 70 species, two species are mainly distributed in the Japanese Islands. *Dugesia japonica* is habitats widely in the Far East (the Japanese Islands, Taiwan, the Korean Peninsula, China, and Primorskiy in Russia) (Kawakatsu et al., 1995). *Dugesia ryukyuensis* is recorded in the Southwest Islands of Japan (Nansei Shoto) and the lowland areas in the Kyushu distinct on the East China Sea (Tamura et al., 1998). Both species have very similar morphologically, however, a karyotype of *D. japonica* (n=8, 2x=16, 3x=24) differs from *D. ryukyuensis* (n=7, 2x=14, 3x=21) in having an asymmetrical penis papilla without a well-developed valve surrounding its basal part, and a well-developed vagina (Kawakatsu et al., 1976).

Nucleotide sequence information of several gene regions on mitochondrial DNA (mtDNA) has been used for evaluating phylogenetic relationships within species and between closely related species because these regions show sufficiently high rates of nucleotide substitution (Avise 1994). The cytochrome oxidase subunit I (*coI*) have been used in phylogenetic relationships within Tricladida (Marta et al., 2008, Lázaro et al., 2009). Partial sequences of *coI* of *D. japonica* were reported in Bessho et al. (1992a, b). Recently, the complete mtDNA sequence of *D. japonica* and most of the mtDNA sequence of *D. ryukyuensis* were determined.

Analysis of the 18S rRNA sequences of species of the family Dugesiidae, Dendrocoelidae and Planaridae showed that members of the family Dugesiidae have two types of 18S rRNA genes, while the rest of the species have only one. The mean sequence divergence value between 18S rRNA type I and the type II sequences is 9% and type II 18S rRNA genes are evolving 2.3 times more rapidly than type I (Salvador et al., 1999). Type II

can be expected as a phylogenetic marker because of its high substitution rate.

In this study, I surveyed detailed genetic population structure in wild populations of *Dugesia* planarians in the Japanese Islands by using molecular phylogenetic analysis based on 18S rRNA type II haplotype and *coI* mitotype sequences. A large number of 18S rRNA type II haplotype and mitotypes were identified. The distribution patterns of each of 18S rRNA type II genotype and *coI* mitotype demonstrated geographical associations. Since comparison of 18S rRNA type II genotypes and *coI* mitotypes among individuals demonstrated particular combinations of genotypes, seven new genetic groups have been defined (R; *D.ryukyuensis*, J-I, -II, -III, -IV, -V and –IV; *D. japonica*).

Materials and Methods

Animals and DNA extraction

The freshwater planarians, *D. japonica* and *D. ryukyuensis*, were collected from 141 different sites in Japan and one in South Korea (Fig. 6). These sites and individuals number are listed in Table 4. OH strain of *D. ryukyuensis* from Keio University was also examined. In addition, two species of the other family, *Seidlia auriculata* from Towada, Aomori, Prefecture, and *Bdellocephala brunnea* from Daisen, Akita, Prefecture were analyzed for outgroup comparison. The phenol-chloroform method was used to isolate genomic DNA from individual organisms after they had been housed without food over a week. Individual organisms were then crushed in 200 ml lysis buffer (0.1 M EDTA [pH 8.0], 0.05 M Tris-HCl [pH 8.0], 0.1 M NaCl, and 1% SDS) with 40 mg Proteinase K (TaKaRa, Shiga, Japan) using a pestle fitting a 1.5-ml Eppendorf tube, and incubated at 55°C for 2 h. An equal volume of phenol/chloroform was added to the tubes. After centrifugation at 15,000 rpm for 5 minutes and transferring the aqueous phase to a new tube, an equal volume of isopropanol acetate was added. After a second centrifugation step, the resultant DNA pellet was washed with 70% ethanol, dried, and suspended in 150 ml Tris-EDTA (TE) buffer (1 mM EDTA, 10 mM Tris [pH 8.0]).

Amplification and sequencing of 18S ribosomal RNA gene type I and II and cytochrome *c* subunit I gene

Two partial 18S rRNA type I regions of *D. japonica* and *D. ryukyuensis* were amplified by using previously published (Salvador C, et al., 1996). Two 18S type I specific primers newly were designed (Table 5). PCRs were performed by using the primer sets 1F and type I_R (anterior half), and type I_F and 9R (posterior half). The thermal cycle profile was 35 cycles of 95°C for 30 s, 58°C for 30 s and 72°C for 1 min.

Two partial 18S rRNA type II regions of *D. japonica* and *D. ryukyuensis* were amplified by using previously published primers. Four 18S rRNA type II specific primers newly were designed (Table 5). PCRs were performed by using the primer sets 1F and type II_3R (anterior half), and type II_2F and 9R (posterior half). Also, these of *D. ryukyuensis* were amplified by using the primer set 1F and type II_4R (anterior half), and type II_4F and 9R (posterior half). The thermal cycle profile was 35 cycles of 95°C for 30 s, 58°C for 30 s and 72°C for 1 min.

The partial *coI* regions of *D. japonica* and *D. ryukyuensis* were amplified by using previously published primers (Bessho et al. 1992 a, b). The PCR was performed by using the primer set CO1_f and CO1_r (Table 5). PCR was done in a 10-μl reaction mixture containing 6.7 μl water, 1.0 μl 5 × Takara EX taq buffer, 0.025 μl EX taq polymerase (TaKaRa, Shiga, Japan), 4 mM dNTP, 2.5 μM of each primer, and 1.0 μl of template. The thermal cycle profile was 35 cycles of 95°C for 30 s, 60°C for 30 s and 72°C for 30 s. 18S rRNA and *coI* of *S. auriculata* and *B. brunnea* were amplified by using the primer sets 1F and 9R, and CO1_f and CO1_r.

Sequencing was performed using the BigDye terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) in a 3130 DNA Analyzer (Applied Biosystems). Sequencing primers are shown in Table 5.

Alignment and phylogenetic analysis

The complete sequences of 18S rRNA type I (accesion number U31084) and type II (accesion number U31085) of *Schmidtea mediterranea* were multiple-aligned by using the MAFFT (Katoh et al. 2009) software using the default option. The Gblocks ver. 0.91 software package (Castresana et al. 2000) was used to exclude ambiguously aligned proportions. The settings for the 46 species dataset were: minimum number for a conserved position, 46; minimum number for a flank position, 15; maximum number of non-conserved positions, 8; minimum length of a block, 10; allowed gap

positions, none. Phylogenetic analyses were performed by using the maximum likelihood (ML) and Bayesian inference (BI) methods. GTR+G+I model selected by Kakusan4 (Tanabe 2007) was used in ML and BI analysis. ML analyses were performed with RAxML v. 7.2.6 (Stamatakis 2006). Five hundred bootstrap replications were performed to infer the support of clades from the tree. BI analysis was performed with MrBayes v. 3.1.2 (Ronquist and Huelsenbeck 2003). The Markov chain Monte Carlo (MCMC) analyses were run with four differentially-heated chains for 8,000,000 generations, with a sampling frequency of 100 generations. Tracer v. 1.5 software (Rambaut and Drummond 2004) was used to graph the log-likelihoods of sampled trees. The point at which the chains became stationary was determined by visual inspection. Earlier trees in the chains were discarded as burn-ins. The remaining sampled trees from each of the two runs were included in the calculations of the Bayesian posterior probabilities. Nucleotide sequence divergences between pairs of 18S rRNA were calculated using the Tamura-Nei model with gamma distribution.

The nucleotide sequence of *coI* of each individual was aligned using ClustalW (J.D. Thompson et al., 1994) implemented in MEGA version 5 (Tamura et al. 2011). Individuals exhibiting identical nucleotide sequences were grouped to generate a sequence haplotype (hereafter called mitotype). Phylogenetic analyses were performed by using the neighbor-joining method (NJ). NJ analyses were performed with MEGA5 software package. The Tamura-Nei model (Tamura and Nei 1993) with gamma distribution was used as the model of nucleotide substitutions. 1000 bootstrap replications were performed to infer the support of clades from the tree. Nucleotide sequence divergences between pairs of mitotypes were calculated using the Tamura-Nei model with gamma distribution.

Results

Genetic diversity and lineage separation of *D. japonica* and *D. ryukyuensis* 18S rRNA type II and *coI* and Polymorphism

The 1745-1768 bp lengths of 18S rRNA typeI and type II was determined, successively, these sequences was aligned and evaluated by Gblocks. The 1690 bp regions of 18S rRNA type II was obtained without insertion or deletion. Across the 141 geographic samples, seventy type II haplotypes were identified and 164 polymorphic sites were found in *D. japonica*. Seven type II haplotypes were identified and 11 polymorphic sites were found in *D. ryukyuensis*. Figure 7 shows the topology based on the BI method. The ML tree had the same topology (data not shown). BI and ML tree of 18S rRNA type I and type II allocated haplotypes of *D. ryukyuensis* into one (clade A), and *D. japonica* into six genetic lineages (clade B-G) (Bayesian posterior probabilities; BPP > 0.90 and bootstrap probabilities; BP > 70%). Monophyly of the ingroups (*D. ryukyuensis* and *D. japonica*) was strongly supported, respectively. ML and BI tree details are shown in the supplementary data because of the large number of taxon (supplementary figure S1). Nucleotide sequence divergences between pairs of typeII in *D. japonica* were 0–0.054 and in *D. ryukyuensis* 0–0.007, and two *Dugesia* species 0.002–0.004. Across the 20 geographic samples, eight type II haplotypes were identified and eight polymorphic sites were found in *D. japonica*. One type II haplotypes was found in *D. ryukyuensis*. Nucleotide sequence divergences of type I in *D. japonica* was 0–0.003 and in *D. ryukyuensis* 0 (Table 6). Those of type I and type II in *D. japonica* were 0.076–0.090 and *D. ryukyuensis* 0.076–0.078.

The 381 bp region of *coI* was aligned for all individuals. *D. ryukyuensis* had three base pair insertions in 283-285 position, and *S. auriculata* and *B. brunnea* had three base pair deletions in 286-288 position. Across the 141 geographic samples, 139 *coI* mitotypes were identified and 173 polymorphic

sites were found in *D. japonica*. Six *coI* mitotypes were identified and 18 polymorphic sites were found in *D. ryukyuensis*. A NJ tree of *coI* allocated mitotypes of *D. ryukyuensis* into two (a1 and a2), and *D. japonica* into 17 genetic lineages (b-q) (BP > 85%) (Fig. 8), Monophyly of the ingroups (*D. ryukyuensis*) was strongly supported, but *D. japonica* was not. The NJ tree details are shown in the supplementary data because of the large number of taxon (supplementary figure S2). Nucleotide sequence divergences between pairs of *coI* in *D. japonica* were 0–0.34, 0.05 in *D. ryukyuensis*, and 0.16–0.24 in two *Dugesia* species (Table 7).

Geographic distribution of 18S rRNA type II haplotypes and *coI* mitotypes

Figure 9 shows the geographic distribution of the 18S rRNA type II haplotypes of the seven genetic lineages making up the BI and ML tree. The distribution patterns demonstrated geographical associations.

The type II haplotypes of clade A was distributed in Okinawa and the lowland areas in the Kyushu district, clade B the western Japan and at two sites in the eastern Japan, clade C at seven sites in Shimane, Hiroshima, Kouchi, Oita, Kumamoto and Nagasaki, clade D the central and eastern Japan and at two sites in Ehime, clade E throughout the Japanese Islands, except of Okinawa and the lowland areas in the Kyushu district, and at one site in South Korea, clade F Kyushu, and clade G at 12 sites in Niigata, Mie, Osaka and Kagoshima Prefecture.

Figure 10 shows the geographic distribution of the mitotypes of the 18 genetic lineages making up the NJ tree. The distribution patterns demonstrated geographical associations.

The mitotypes of group a1 was distributed in Okinawa, group a2 the lowland areas in the Kyushu, b western Japan and at two sites in the eastern Japan, group c1 throughout the Japanese Islands, group c2 the eastern Japan, group d two sites Nagasaki, group e at two sites in Tsushima Island, group f

and j at one site in South Korea, group g the central and western Japan, group h at seven sites in Niigata, Fukushima, Yamagata and Iwate, group i at seven sites in Wakayama, Hyogo, Shimane, Ehime, Kochi and Nagasaki, group k at five sites in Oita, Kumamoto, Kagoshima, Ehime, Kochi and Nagasaki, group l at seven sites in Shimane, Hiroshima, Kouchi, Oita, Kumamoto and Nagasaki, group m at two sites in Mie and Osaka, group n at four sites in Kagoshima, group o at eleven sites in Niigata and Yamagata, group p at three sites in Aichi and Ehime, group q the central and eastern Japan.

Discussion

While *coI* of *D. japonica* had a very large number of variations, the phylogenetic relationships among genetic groups (c-q) become unclear. Using only the first and second codon positions for the *coI* fragments to try to avoid saturation problems for this molecule, however, the problem was not resolved (date not shown). By determining longer *coI* sequences or analyzing other mtDNA genes, it is expected that phylogenetic relationships among genetic groups (c-q) become clear.

Genetic divergences between 18S rRNA type I and type II was 0.076–0.090 about as similar values as previously (Salvador et al. 1999). The BI and ML trees show one clade linage in *D. ryukyuensis* and six in *D. japonica*. The evolutionary rate of 18S rRNA type II sequence was estimated to be a 7.5×10^{-4} /site/Myr in the family Dugesiidae (Salvador et al. 1999). On the basis of this rate and percentage differences, the divergence time between *D. ryukyuensis* and *D. japonica* is estimated at 24.7-36 million years ago (mya), clade B and other at 20.7-25.3 mya, clade C at 9.3-20 mya, clade D at 10-18.7 mya, clade E at 5.3-16 mya , F and G at 6.0-12.7 mya.

The geographical distribution of type II and mitotype was a similarity was observed. Comparison of 18S rRNA type II genotypes and *coI* mitotypes among individuals demonstrated particular combinations of genotypes (Table 8). I defined seven new genetic groups (R, J-I, -II, -III, -IV, -V and -IV) based on 18SrRNA type II genotypes and *coI* mitotypes (Fig. 11). The distribution pattern of J-I, -III and -V demonstrated geographical associations, but J-II, -IV, and -VI was not. J-II, -IV and -VI showed that 18S rRNA typeII sequences were similar each ingroup, but *coI* were geographically differentiated. Thus, they had some kind of event to divide the habitat in the past rather than a simple artificial movement. J-I showed that 18S rRNA typeII sequences were similar a wide area including Korea,

but *coI* were geographically differentiated. It is considered the expansion of the distribution is assumed to be a newly event.

Geographic distributions of J-I and -II are over in the East Japan, those of J-I, -III and –IV in the West Japan, and those of J-IV and -V the Kyushu distinct. According to Avise (2000), Category II shows deep gene tree and major lineages broadly sympatric. Avise suggests that this pattern is the result of either a species with a large effective population size and high levels of gene flow or, more commonly, secondary contact between divergent lineages. The particular combinations of genotypes were also demonstrated in the populations where were observed two or more genotypes (site 12, 21, 27, 43, 47, 50, 51, 64, 71, 76, 77, 79, 81, 82, 85, 90, 91, 92, 93, 101, 103, 110, 121, 128, 129, 137 and 138; Table 4). These results suggested that there is genetic isolation among J-I, -II, -III, -IV, -V and -VI, and existence of six sibling species in *D. japonica*. Reproductive isolating mechanism in that can be solved by clarifying the model of reproduction of population where were observed two or more genotypes. Meanwhile, individuals in Kagoshim (site 134) had 18S rRNA type II genotype (clade A) from *D. ryukyuensis* and *coI* mitotype (group k) from *D. japonica* (3/9). This suggests that hybridization have occurred between *D. ryukyuensis* and *D. japonica* in the past.

Tamura et al (1991) hypothesized that *D. japonica* invaded the Southwest Islands after the early Quarternary period, probably through both a southern route (via China and Taiwan to the Old Sakishima Islands, the Old Okinawa Islands, and the Old Amami Islands) and a southeastern route (via China and Taiwan to the Korean Peninsula, the Old Kyushu Island, and the Old Amami Islands). Our results showed that divergence times among six groups of *D. japonica* and genetic those distribution, respectively. However, the estimated divergence time of genetic groups was older. Theses indicate that genetic six groups of *D. japonica* migrated into the Japanese Islands by the independent event up to six times and two or more route rather than

differentiated into six groups after the migration.

Conclusion

D. japaonica was a high level genetic diversity for mitochondrial and nuclear DNA. I determined the complete mtDNA sequence of *Dugesia japonica* and most of the mtDNA sequence of *Dugesia ryukyuensis*. The gene order of the mitochondrial genome from the *Dugesia* species showed no clear homology with either the Neodermata or other free-living Rhabditophora. This indicates that the platyhelminths exhibit a great variability in mitochondrial gene order. This is the first complete sequence analysis of the mitochondrial genome of a free-living Rhabditophora. As more mitochondrial genomes are characterized for a wider taxonomic coverage of the phylum Platyhelminthes, the phylogenetic and taxonomic utility of these features can be evaluated. A large number of 18S rRNA type II haplotype and mitotypes were identified in *Dugesia* species in the Japanese Islands. Seven new genetic groups (R; *D. ryukyuensis*, J-I and –IV; *D. japonica*) were defined. Furthermore, the particular combinations of genotypes were also demonstrated in the populations where were observed two or more genotypes. This suggested that there is genetic isolation among J-I, -II, -III, -IV, -V and -VI, and existence of six sibling species in *D. japonica*. Reproductive isolating mechanism in *D. japonica* can be solved by clarifying the model of reproduction of population where were observed two or more genotypes.

Figure legends

Fig. 1. Arrangement of the mitochondrial genomes of *Dugesia japonica* and *D. ryukyuensis*. The protein-coding and rRNA genes are shown as open arrows, tRNAs as lines, and non-coding regions as open boxes. *atp6* refers to ATP synthase; *coI, II, III* refer to the cytochrome oxidase subunits; *cytb* refers to cytochrome *b*; and *nad1–6* refers to the nicotinamide dehydrogenase subunits. tRNAs are denoted by single letters. The two long non-coding regions are denoted as NC1 and NC2.

Fig. 2. Predicted secondary structures of the putative tRNA genes in *D. japonica*. Sequences are illustrated from the 5′ to the 3′ end.

Fig. 3. Predicted secondary structures for the (A) small and (B) large subunits of the rRNA gene in the mitochondrial genome of *D. japonica*. Bonds between C:G and U:A are indicated by straight lines, those between U:G by small closed circles, those between A:C by large closed circles, and those between A:G by open circles.

Fig. 4. Putative secondary structures of the NC2-J in the mitochondrial genome of *D. japonica*.

Fig. 5. Comparison of gene order. The gene order of *D. japonica* and *D. ryukyuensis* was compared to the partial sequences of *M. lineare* (Microstomidae), as well as to the complete sequences of *Trichobilharzia regent* (Trematoda), *Gyrodactylus salaris* (Monogenea), and *Hymenolepis diminuta* (Cestoda).

figure legend

Fig. 6. Collection sites of *D. japonica* and *D. ryukyuensis*. The numbers refer to the locations listed in Table 4.

Fig. 7. Tree constructed from a Bayesian and Maximum likelihood analysis of 1690 bp from 18S rRNA type I and type II gene. Numbers around the branches indicate the Bayesian posterior probabilities (>0.90) and bootstrap probabilities (>70%).

Fig. 8. Tree constructed from a Neighbor-joinig analysis of 381 bp from a

partial *coI* gene. The sequence divergences were calculated with Tamura-Nei model (Tamura and Nei 1993) with gamma distribution used as the model of nucleotide substitutions. 1000 bootstrap replications were performed to infer the support of clades from the tree. Numbers around the branches indicate bootstrap values (>70%).

Fig. 9. Geographic distribution of 18S rRNA type II haplotypes. Clustering of type II haplotypes was inferred from the BI and ML tree (Fig. 7). (A) Geographic distributions of clade A and E, and (B) those of clade B, C, D, F and G.

Fig. 10. Geographic distribution of mitotypes. Clustering of mitotypes was inferred from the BI and ML tree (Fig. 8). (A) Geographic distributions of group a1, a2, c1, c2, d, e, h, i, o and g, and (B) those of group b, f, j, k, l, m, n, p and q.

Fig. 11. Geographic distribution of genetic groups based on 18S rRNA typeII haplotypes and *coI* mitotypes. (A) Geographic distributions of genetic groups R and J-I, (B) those of J-II and -IV, (C) those of J-III, -V and -VI.

Supplementary figure 1. Detail tree constructed from a Bayesian and Maximum likelihood analysis of 1690 bp from 18S rRNA type I and type II gene. Numbers around the branches indicate the Bayesian posterior probabilities (>0.90) and bootstrap probabilities (>70%).

Supplementary figure 2. Detail tree constructed from a Neighbor-joinig analysis of 381 bp from a partial *coI* gene. The sequence divergences were calculated with Tamura-Nei model with gamma distribution used as the model of nucleotide substitutions. 1000 bootstrap replications were performed to infer the support of clades from the tree. Numbers around the branches indicate bootstrap values (>70%).

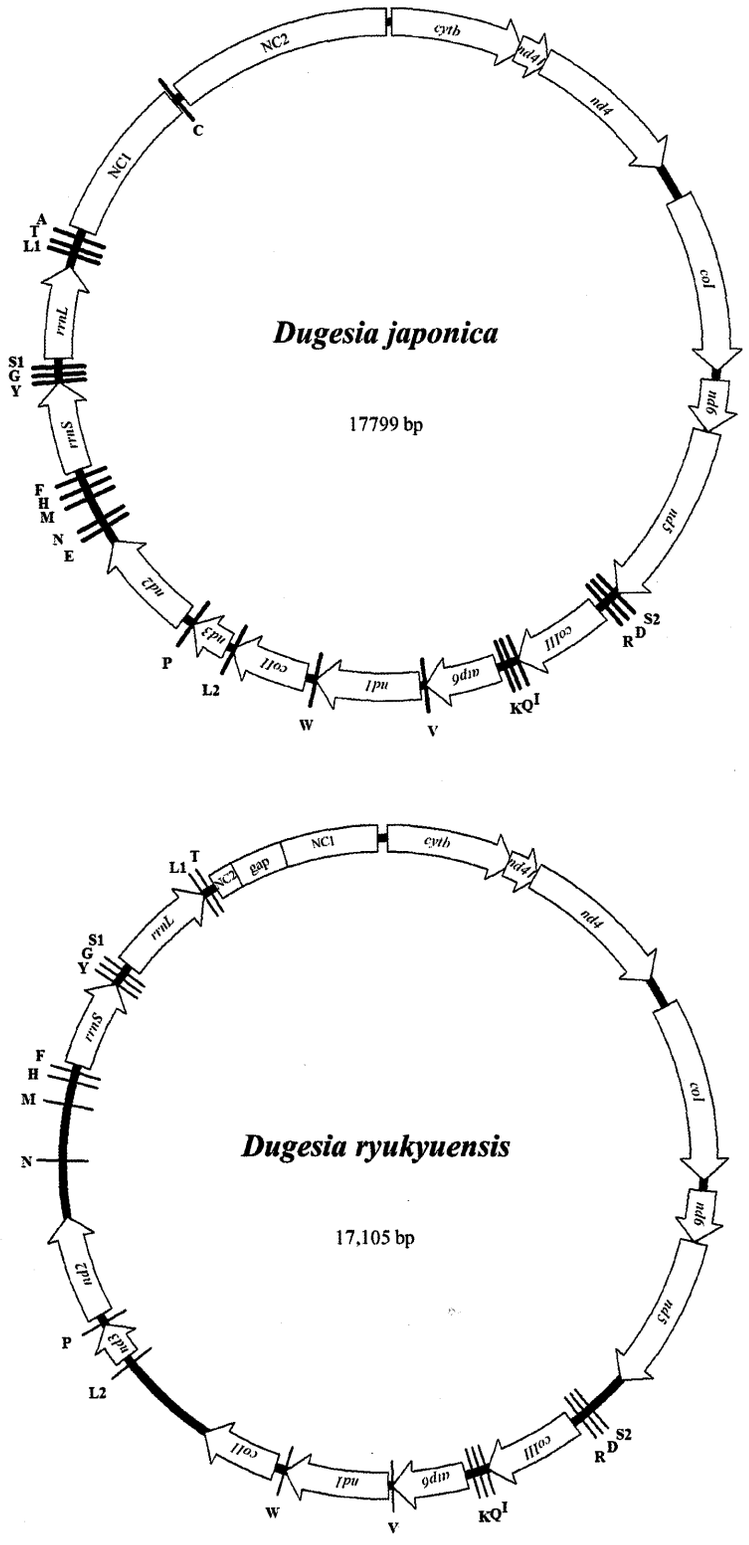


Fig. 1. Arrangement of the mitochondrial genomes of *Dugesia japonica* and *D. ryukyuensis*.

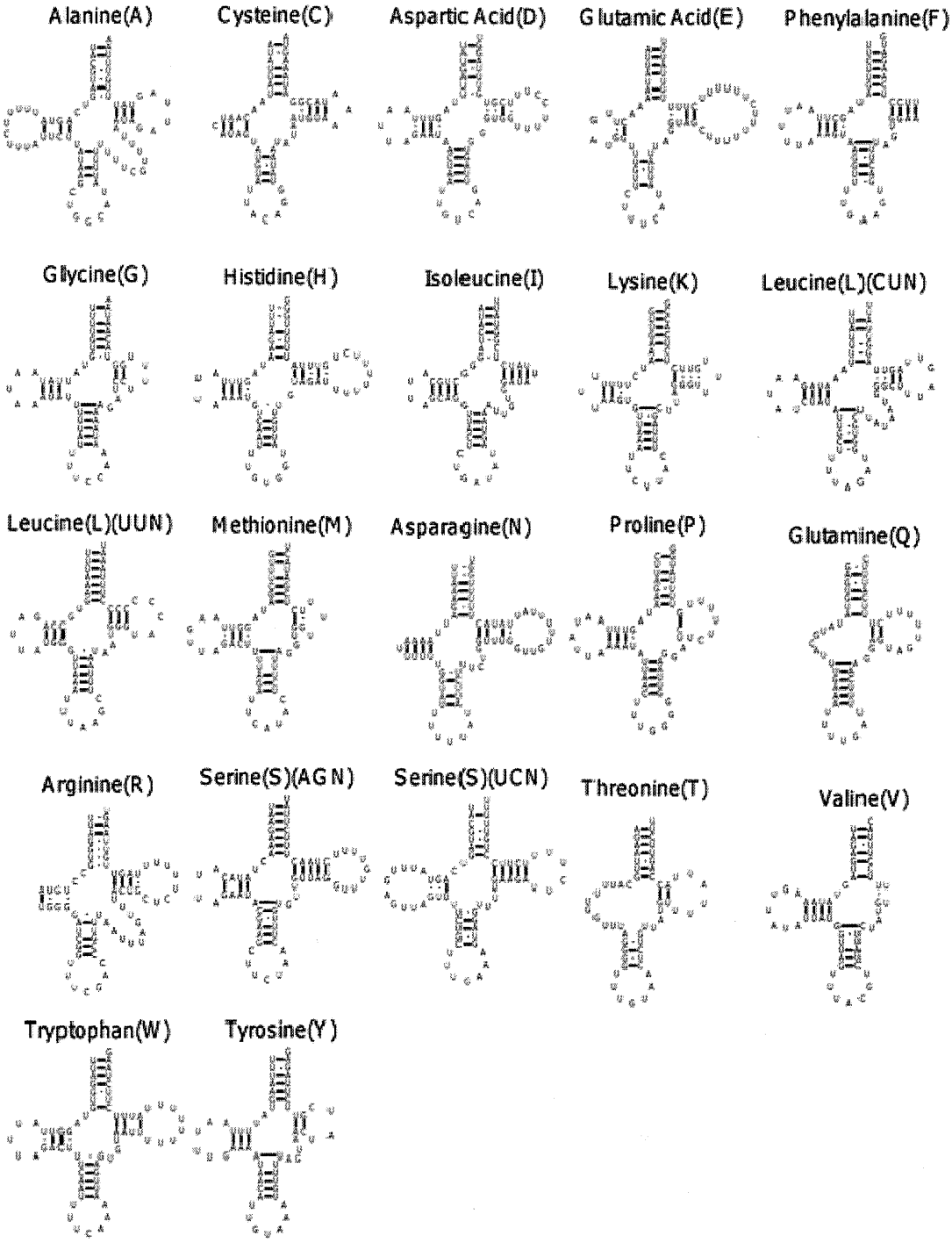


Fig. 2. Predicted secondary structures of the putative tRNA genes in *D. japonica*.

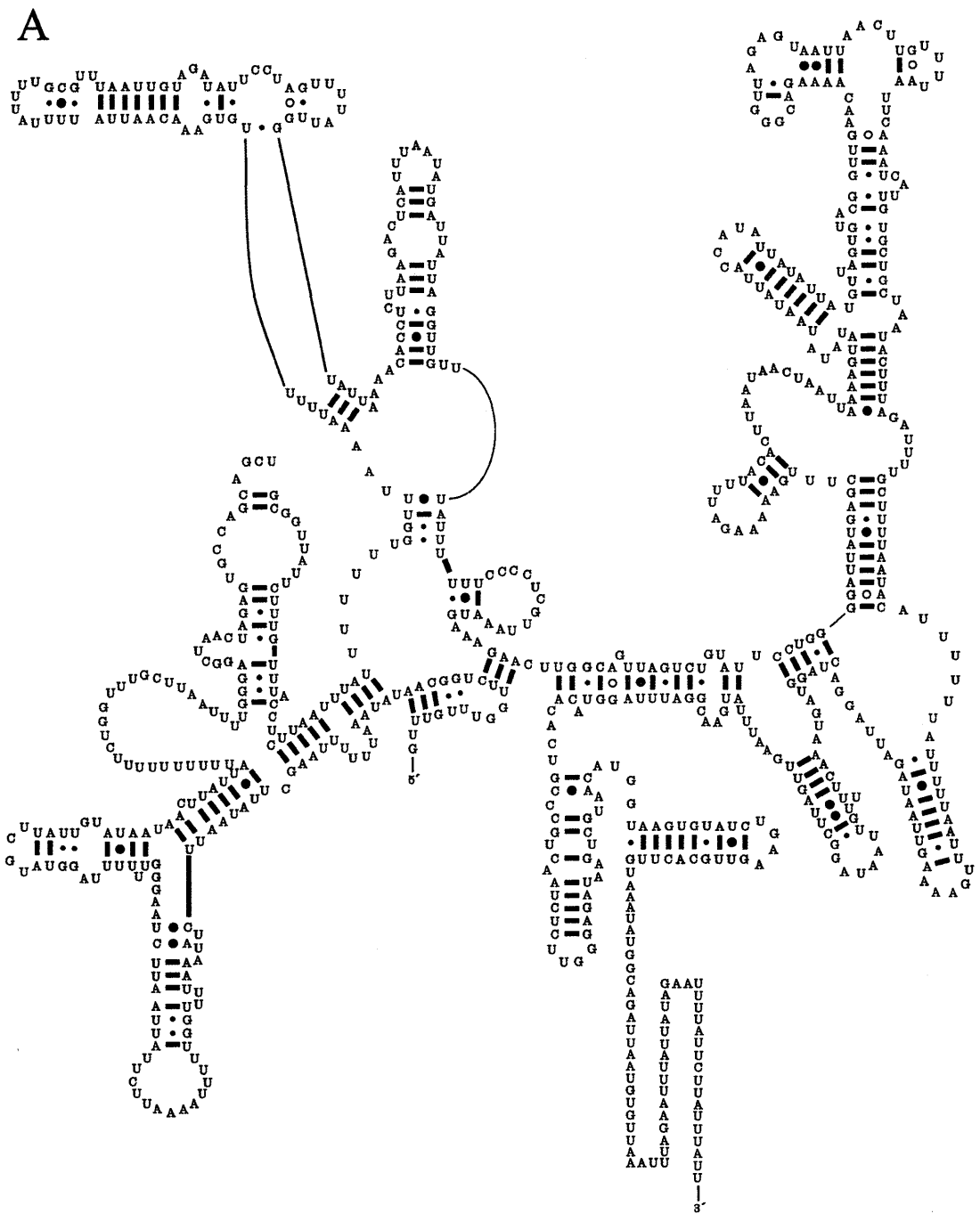


Fig. 3(A). Predicted secondary structures for the small of the rRNA gene in the mitochondrial genome of *D. japonica*.

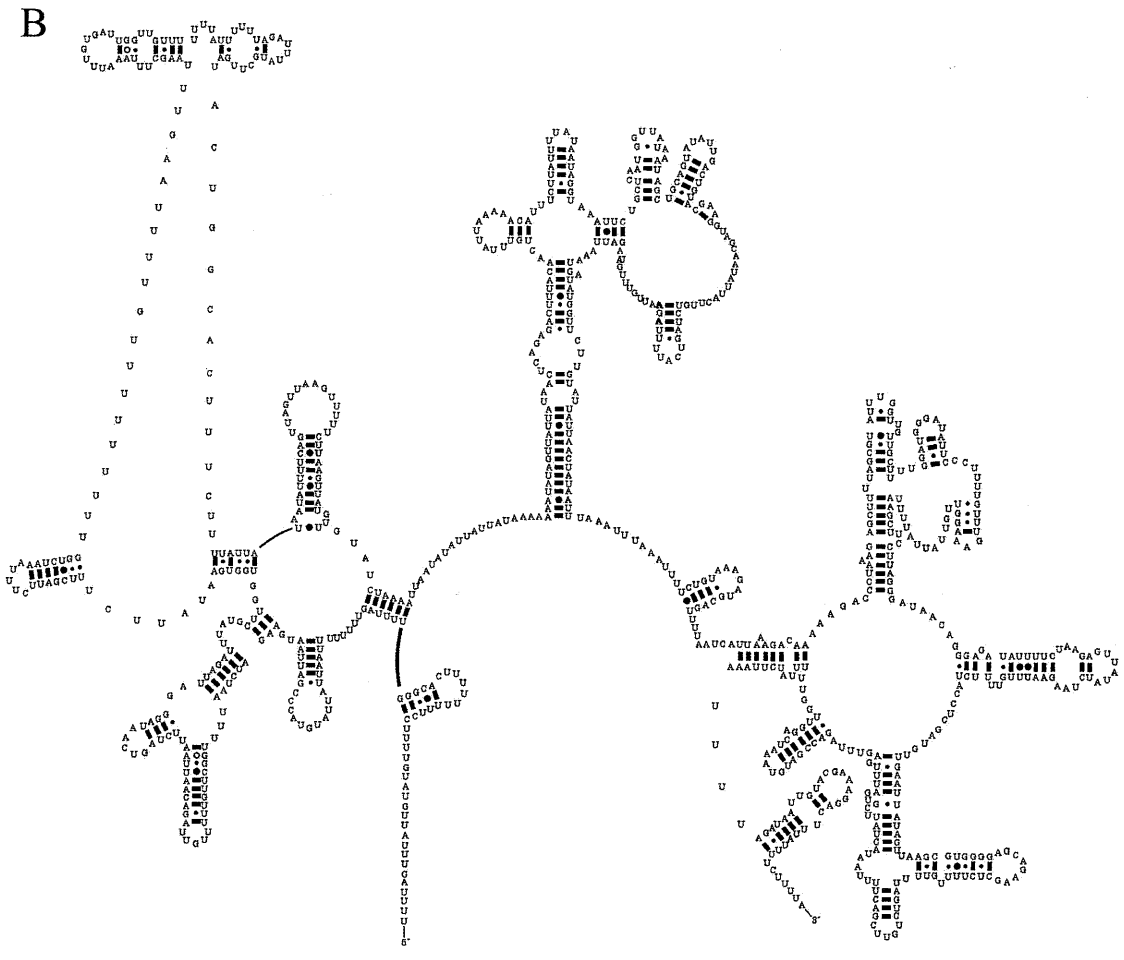


Fig. 3(B). Predicted secondary structures for the large of the rRNA gene in the mitochondrial genome of *D. japonica*.

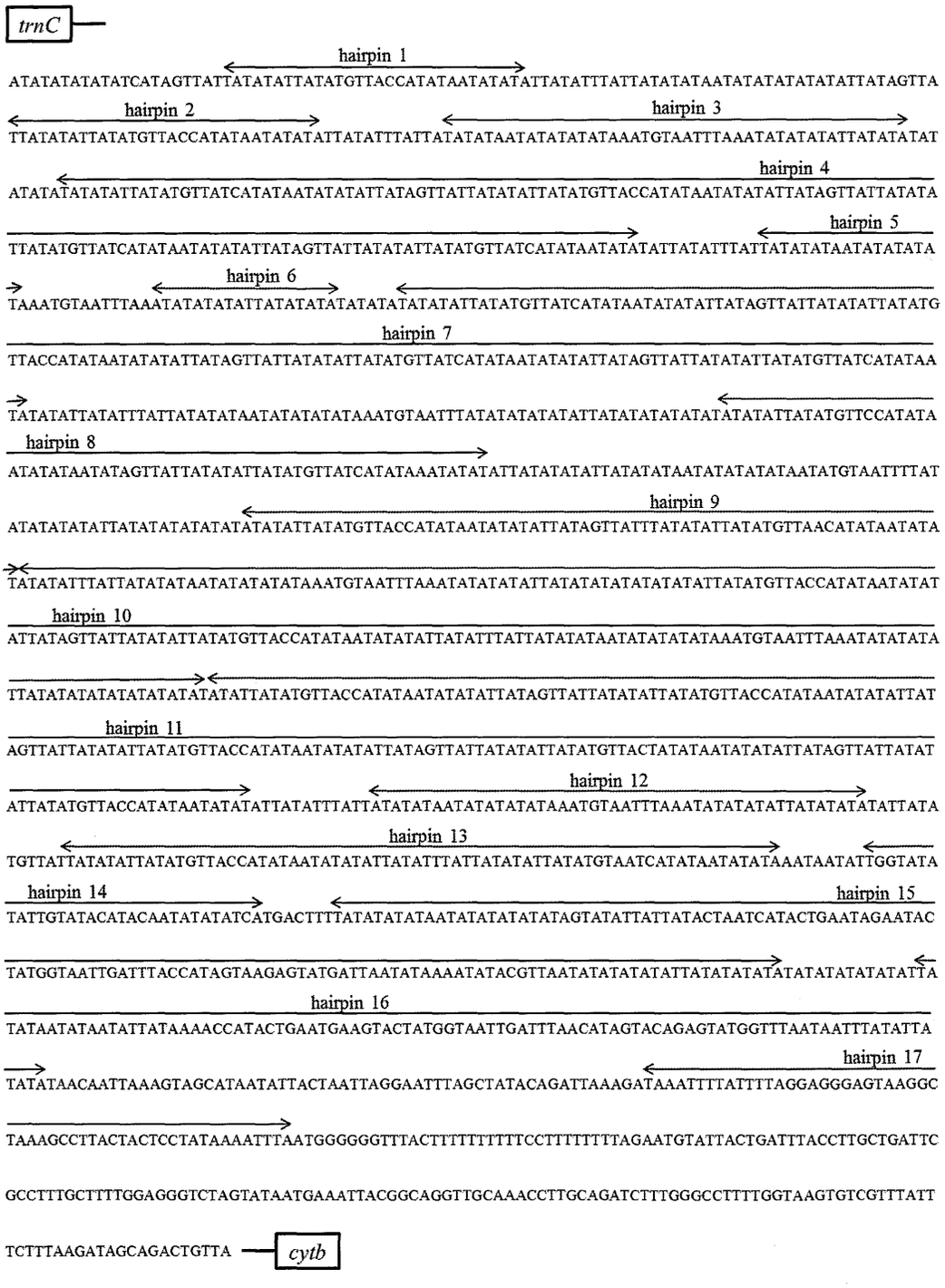


Fig. 4. Putative secondary structures of the NC2-J in the mitochondrial genome of *D. japonica*.

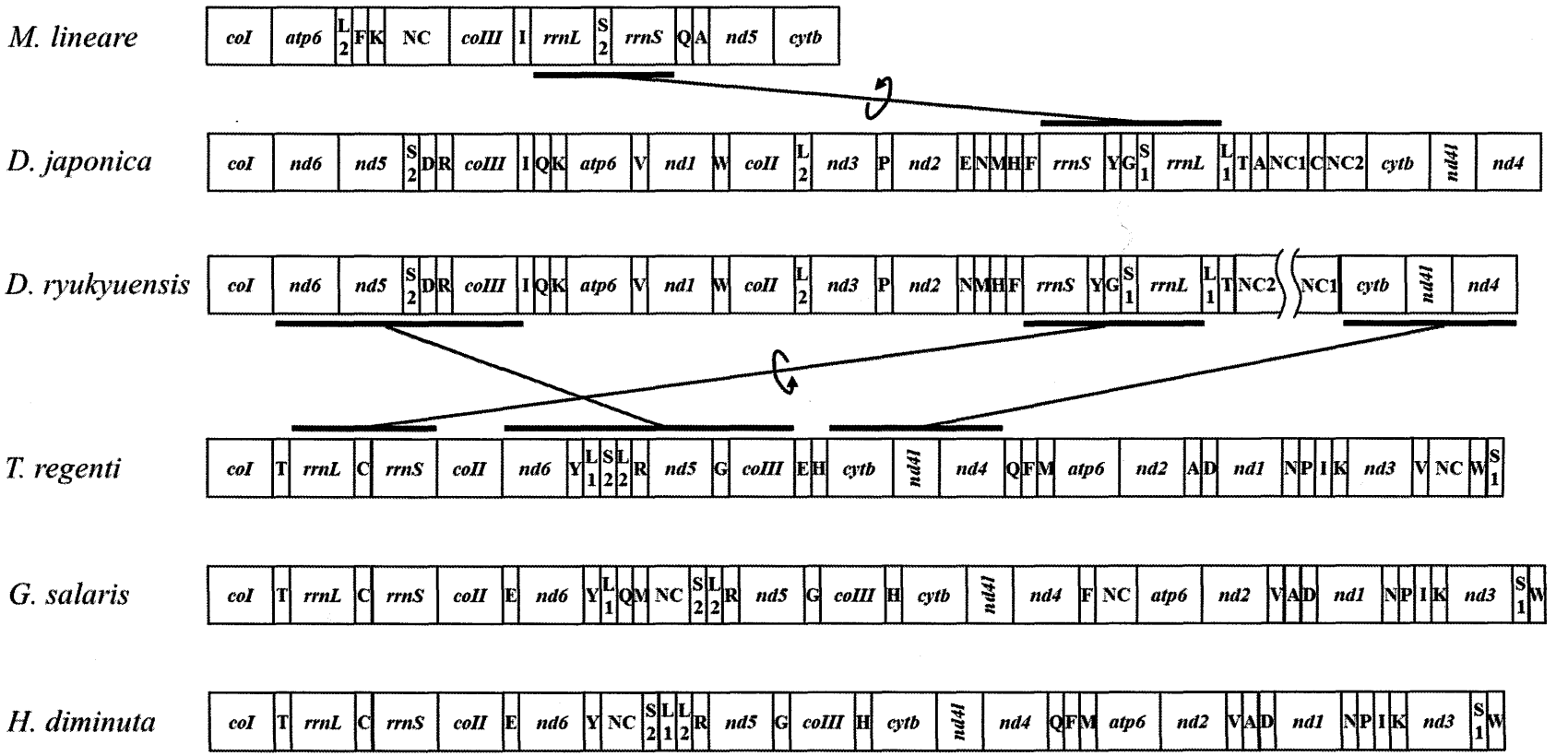


Fig. 5. Comparison of gene order.

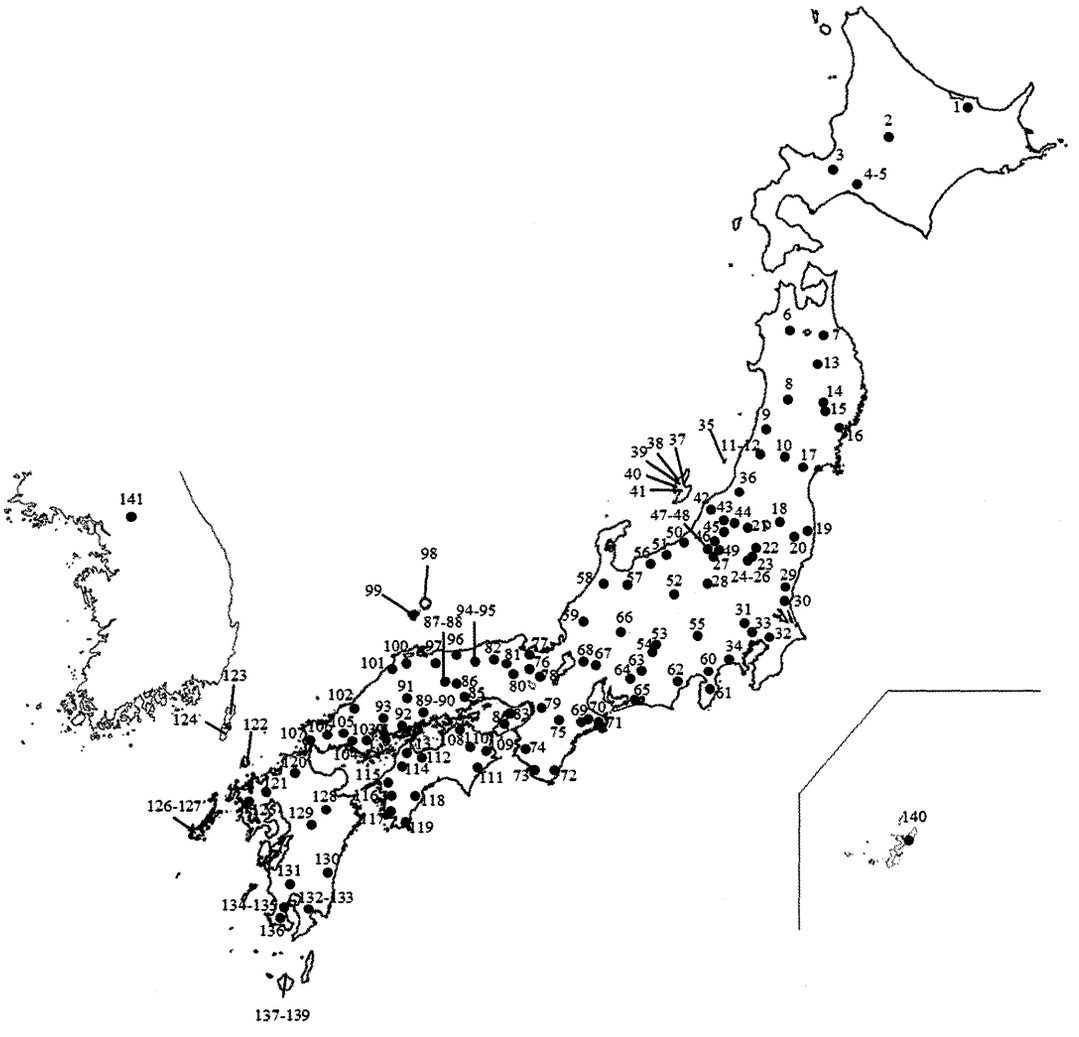


Fig. 6. Collection sites of *D. japonica* and *D. ryukyuensis*.

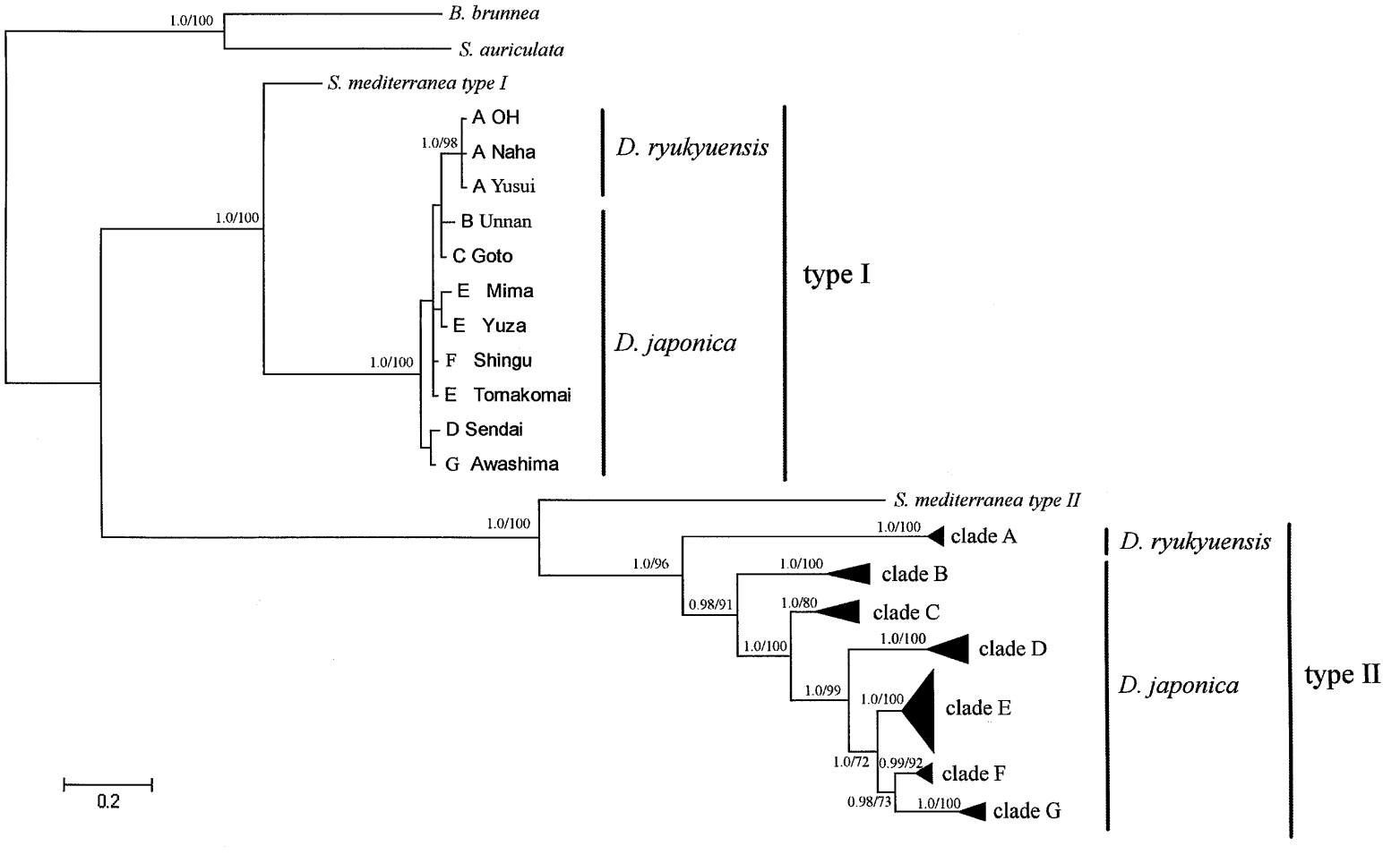


Fig. 7. Tree constructed from a Bayesian and Maximum likelihood analysis of 1690 bp from 18S rRNA type I and type II gene.

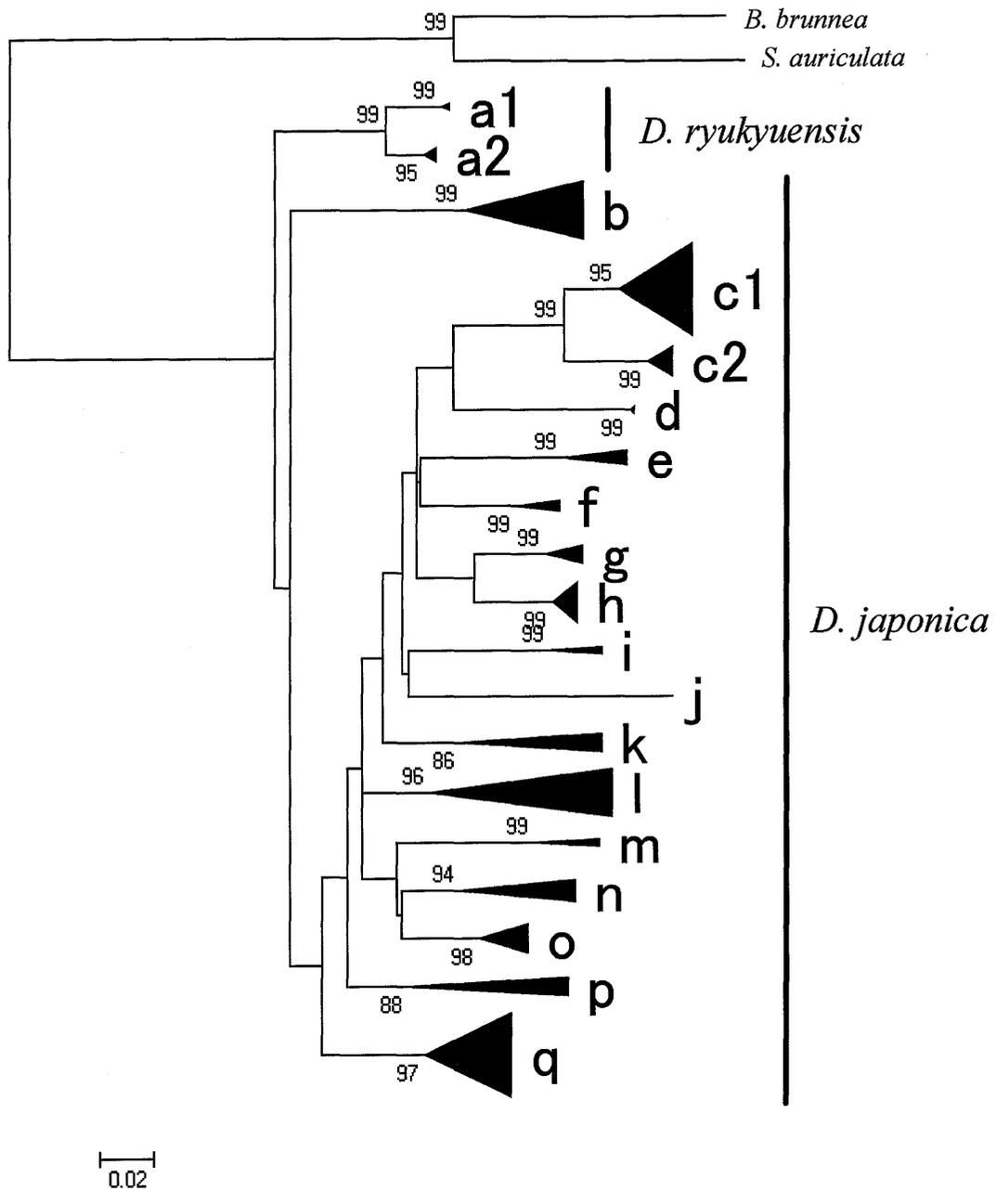


Fig. 8. Tree constructed from a Neighbor-joinig analysis of 381 bp from a partial *coI* gene.

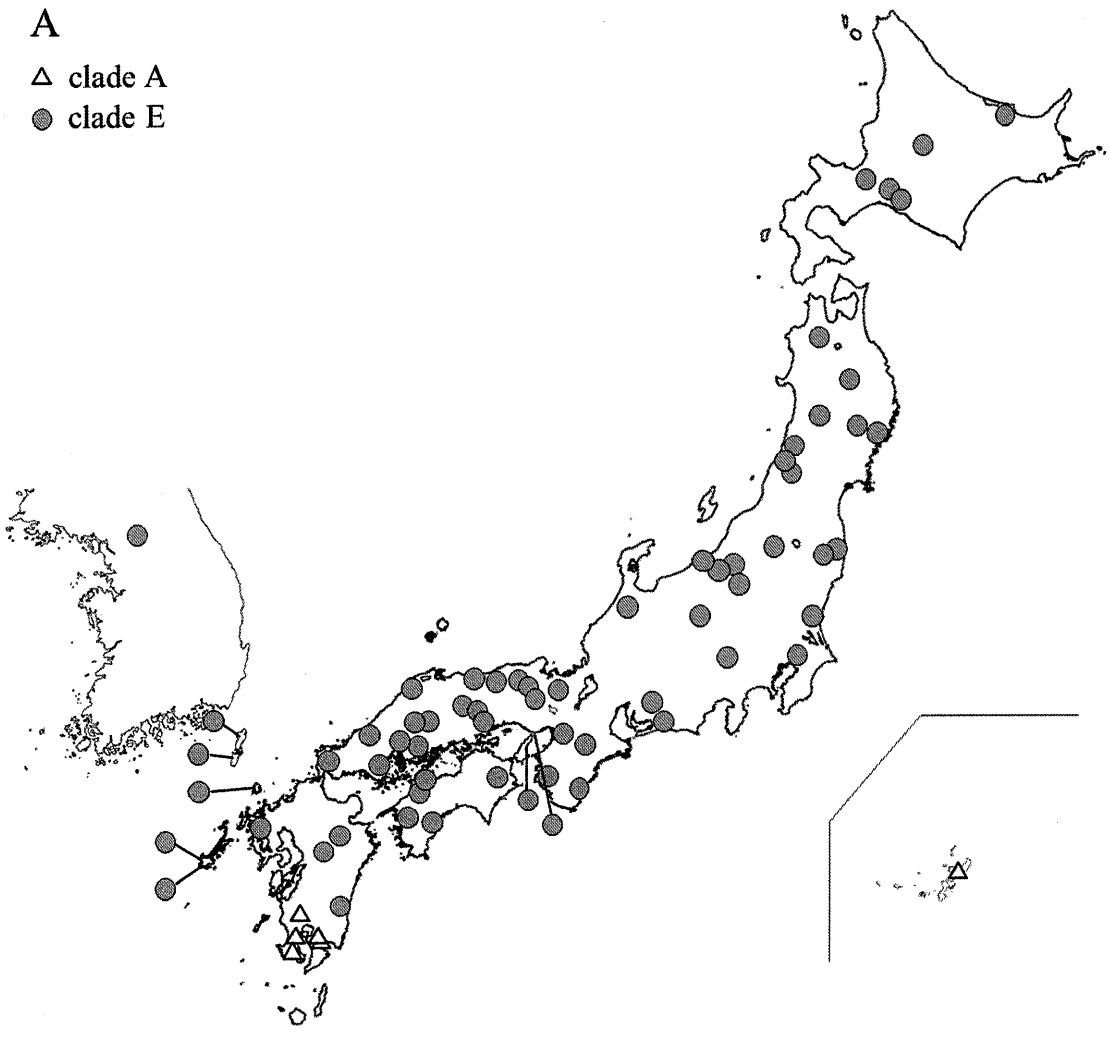


Fig. 9(A). Geographic distributions of clade A and E.

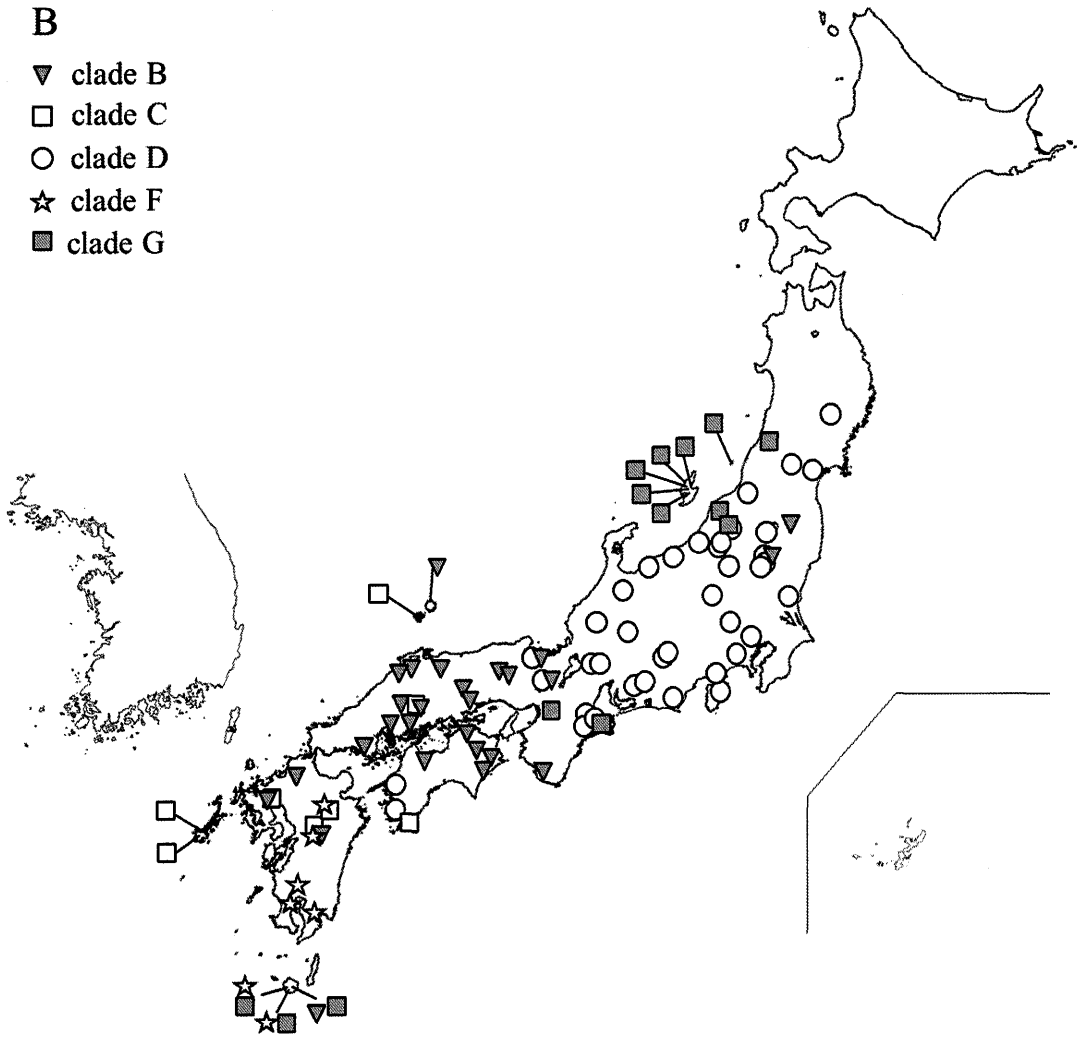


Fig. 9(B). Geographic distributions of clade B, C, D, F and G.

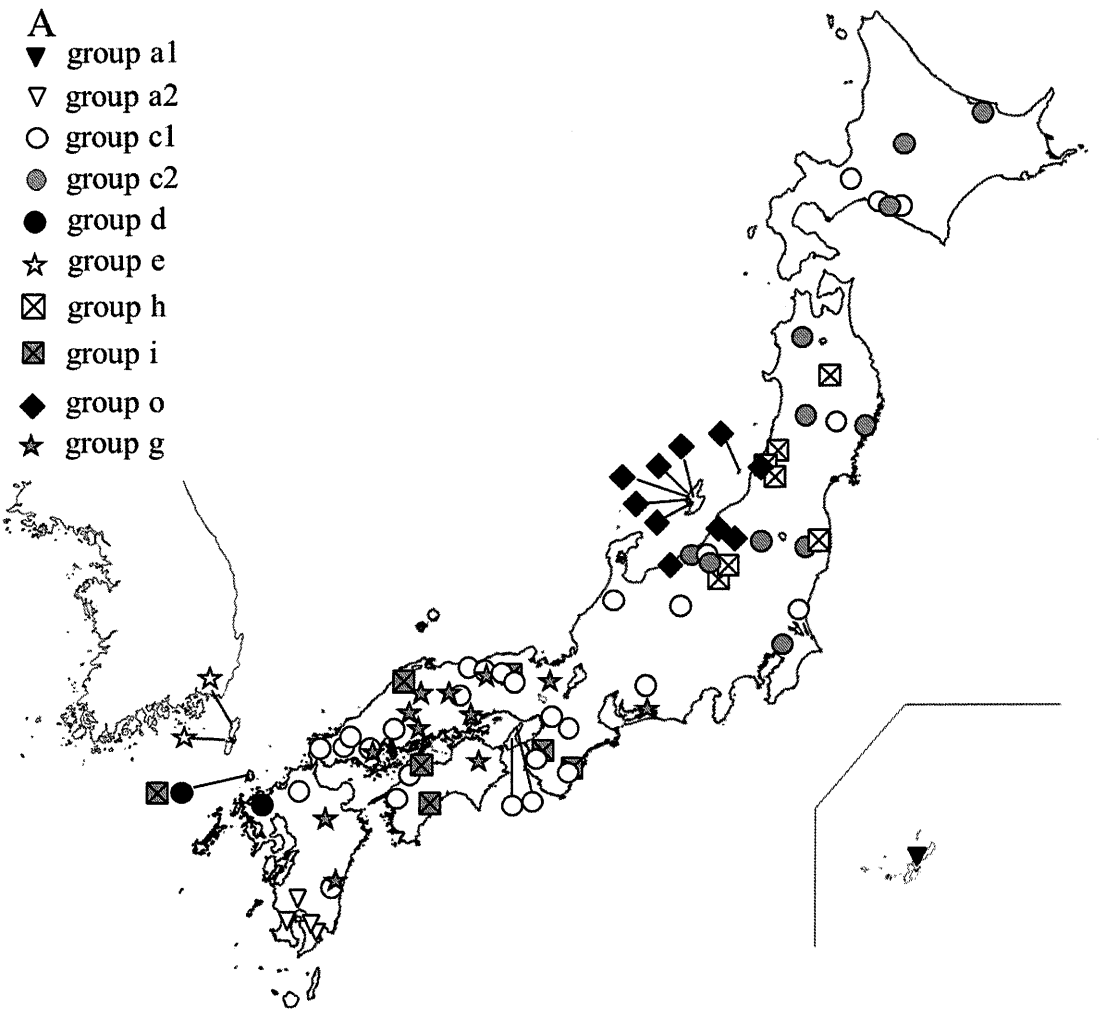


Fig. 10(A). Geographic distributions of group a1, a2, c1, c2, d, e, h, i, o and g.

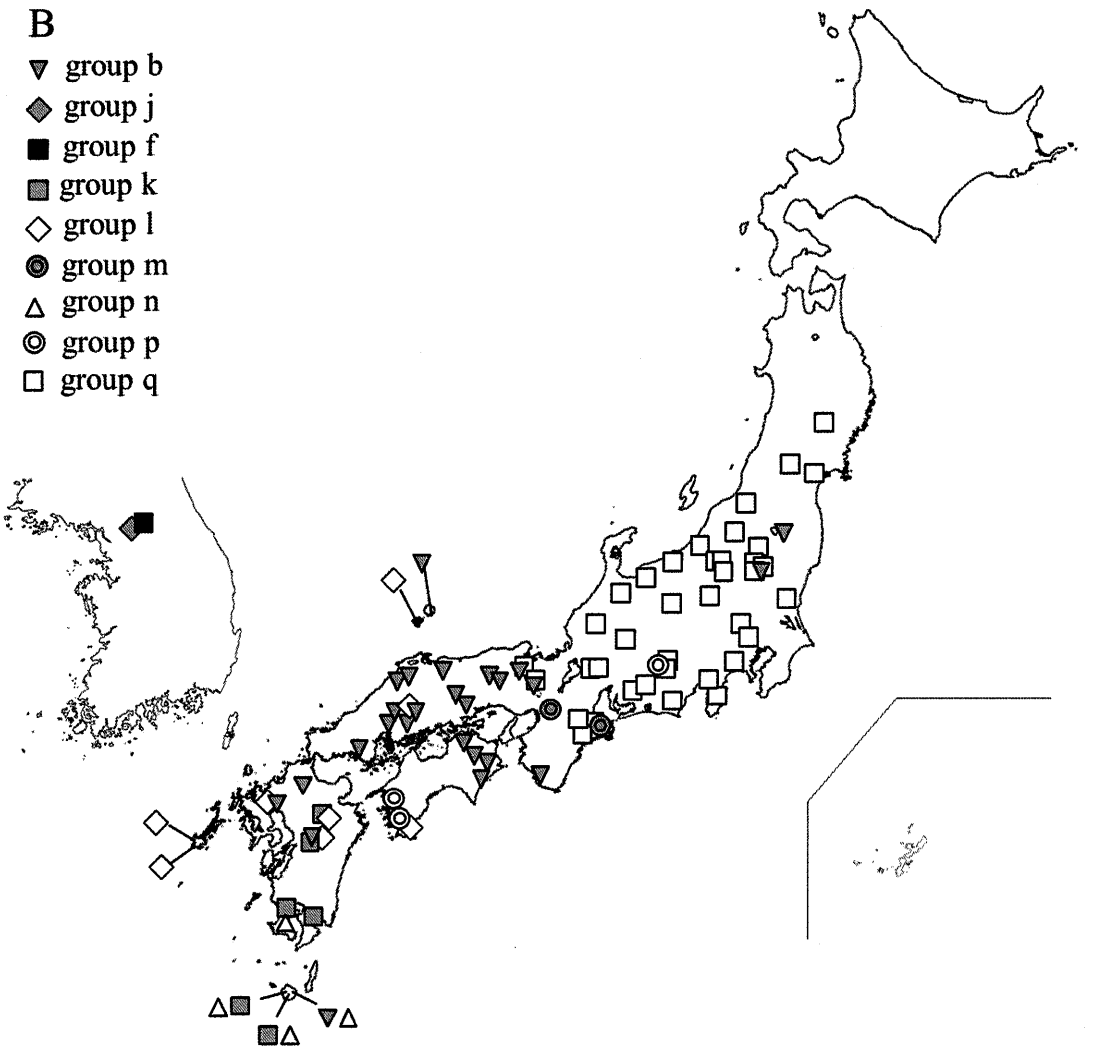


Fig. 10(B). Geographic distributions of group b, j, k, l, m, n, p and q.

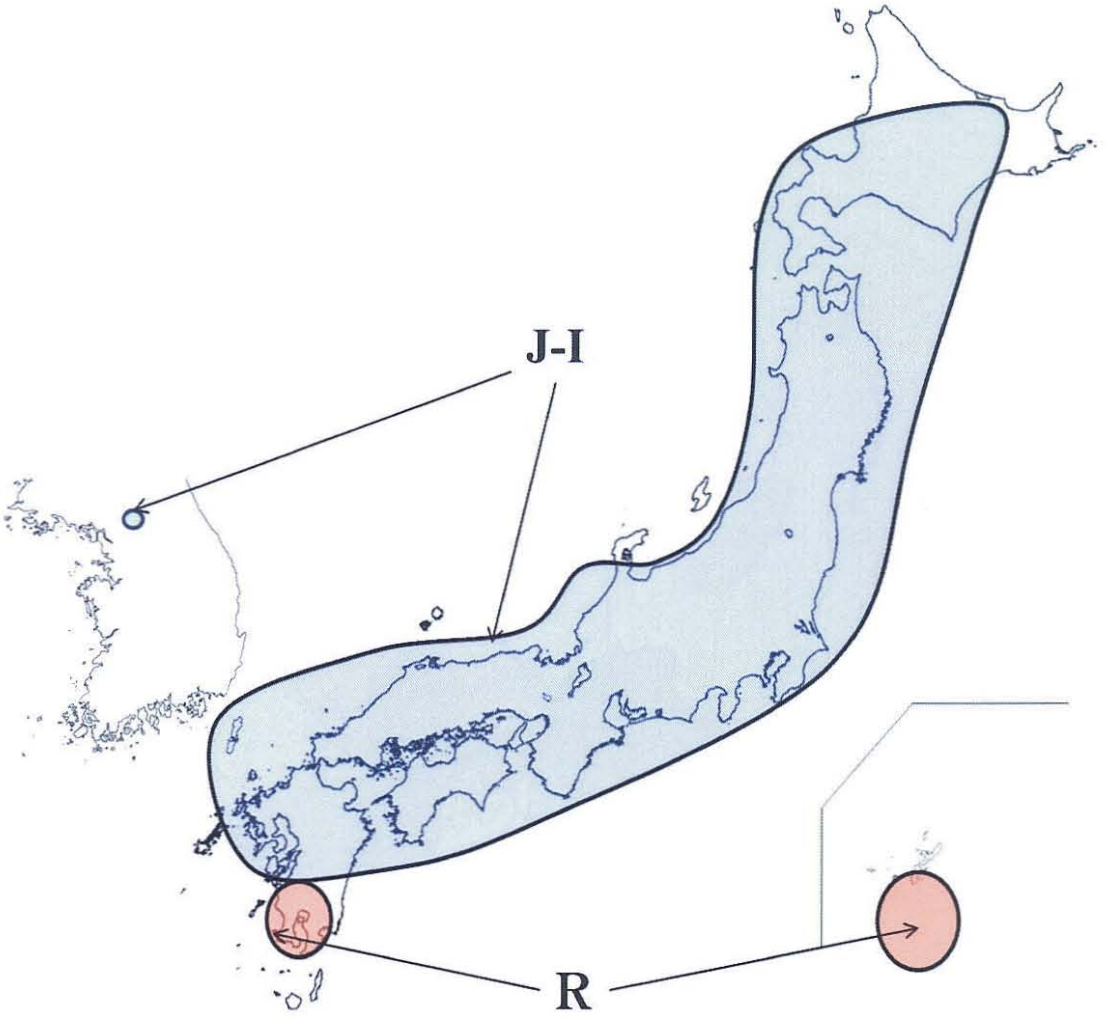


Fig. 11(A). Geographic distributions of genetic groups R and J-I.

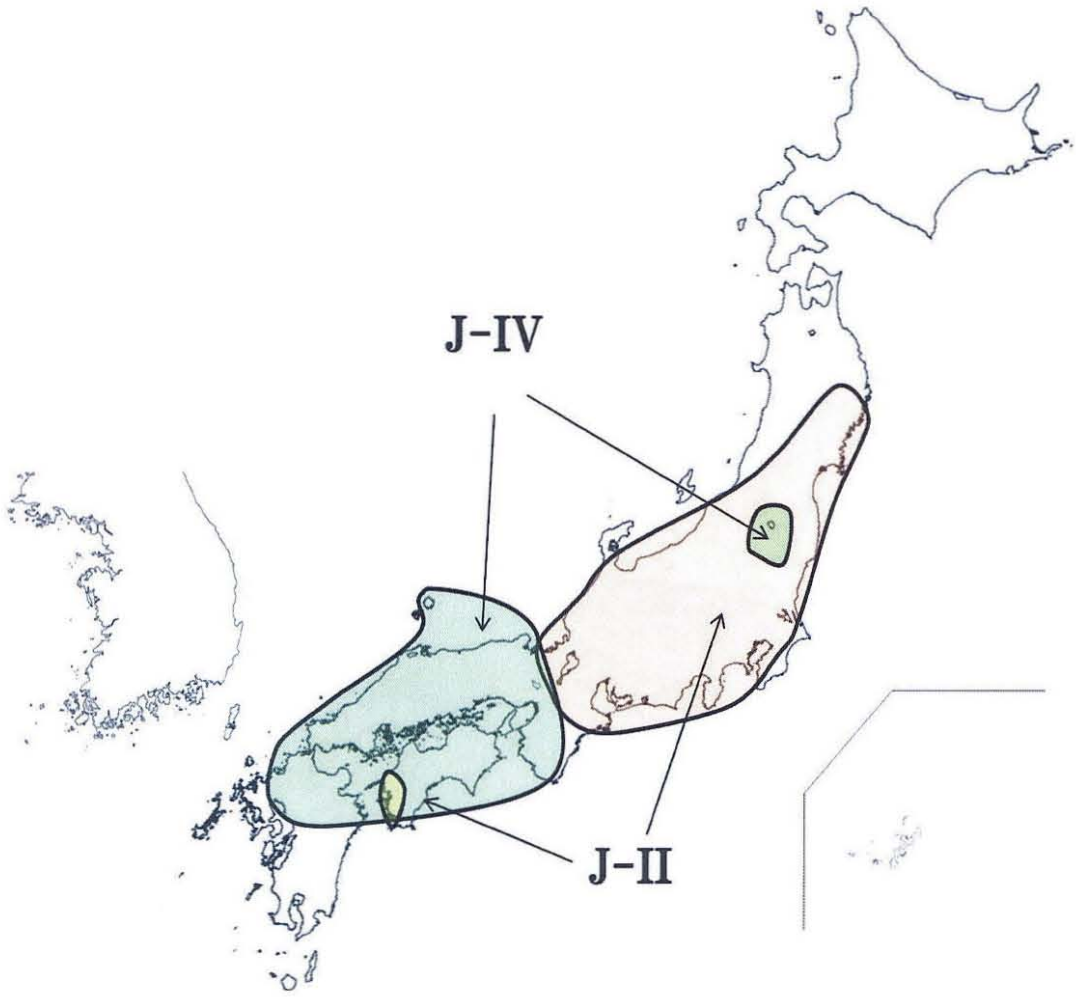


Fig. 11(B). Geographic distributions of genetic groups J-II and –IV.

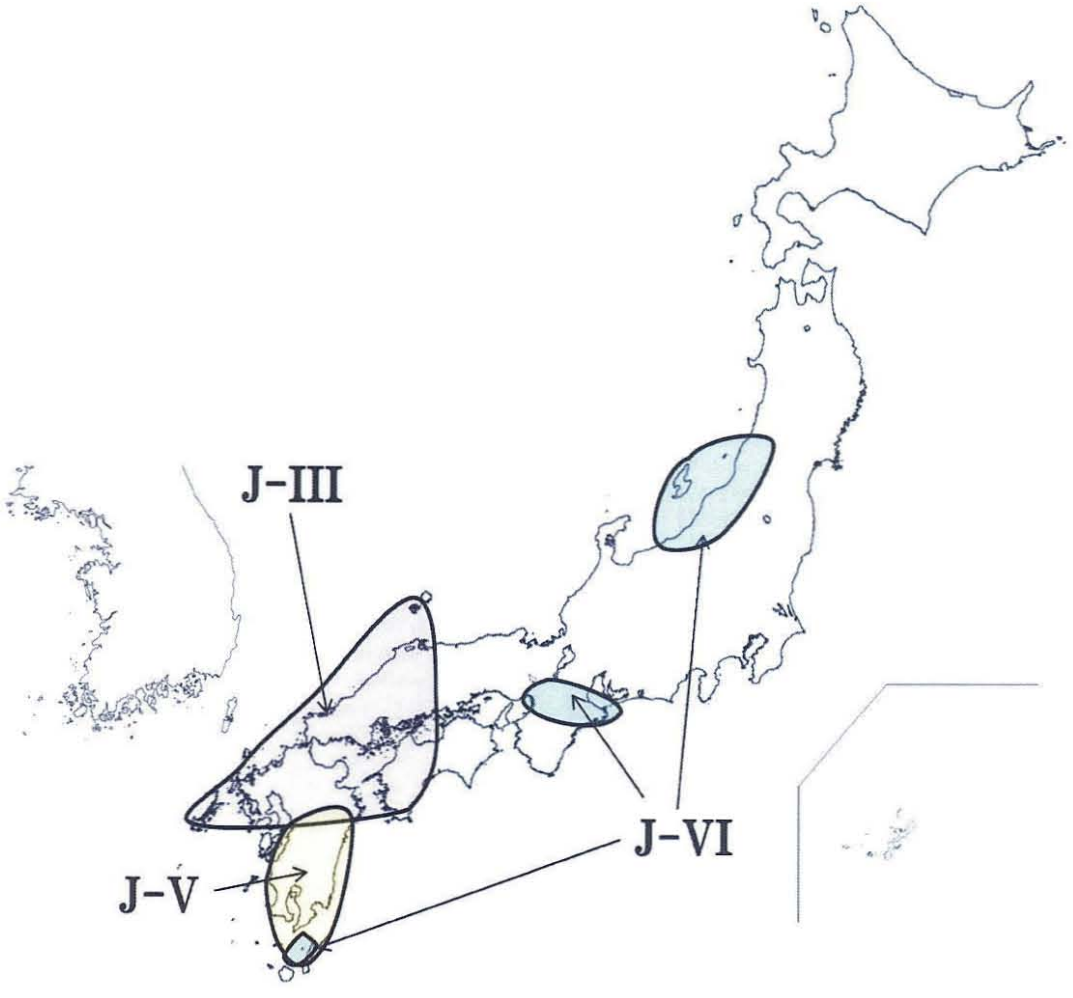
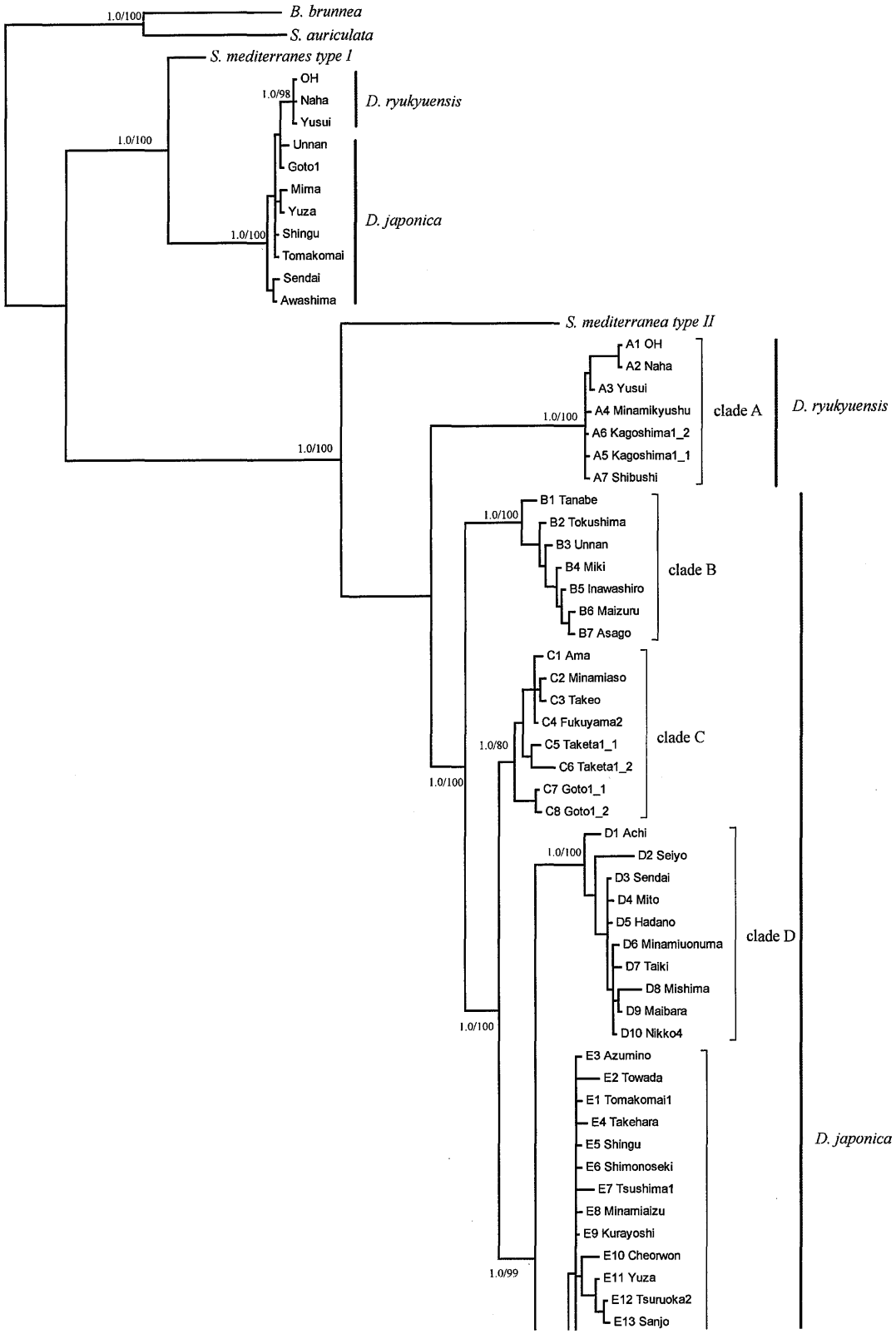
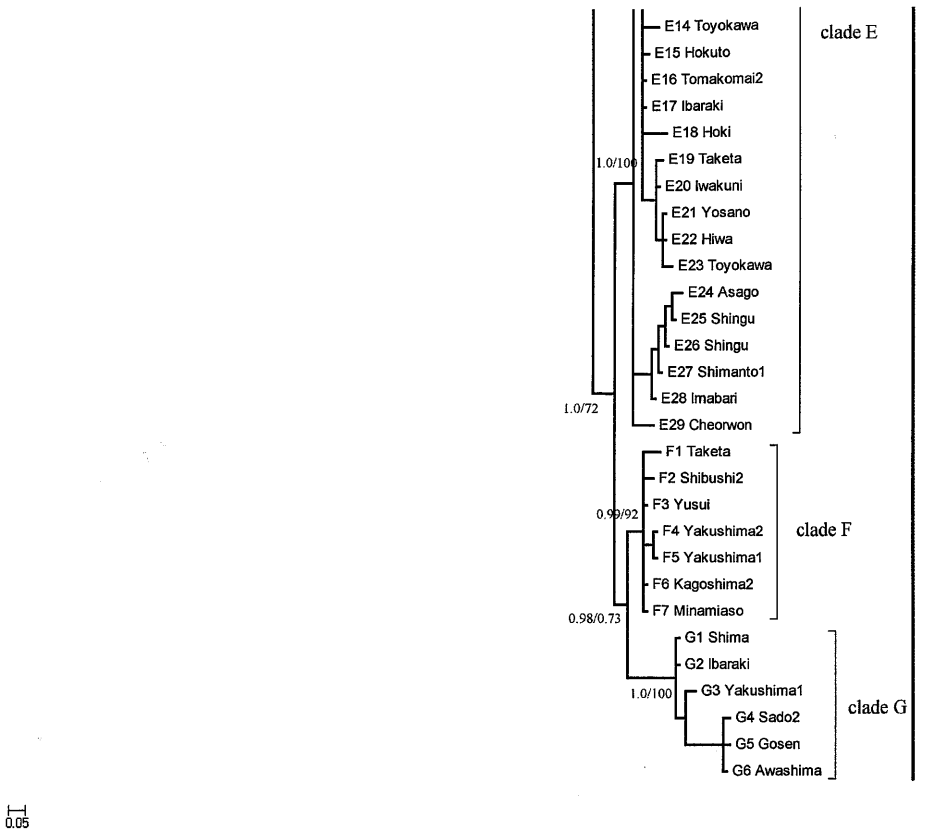
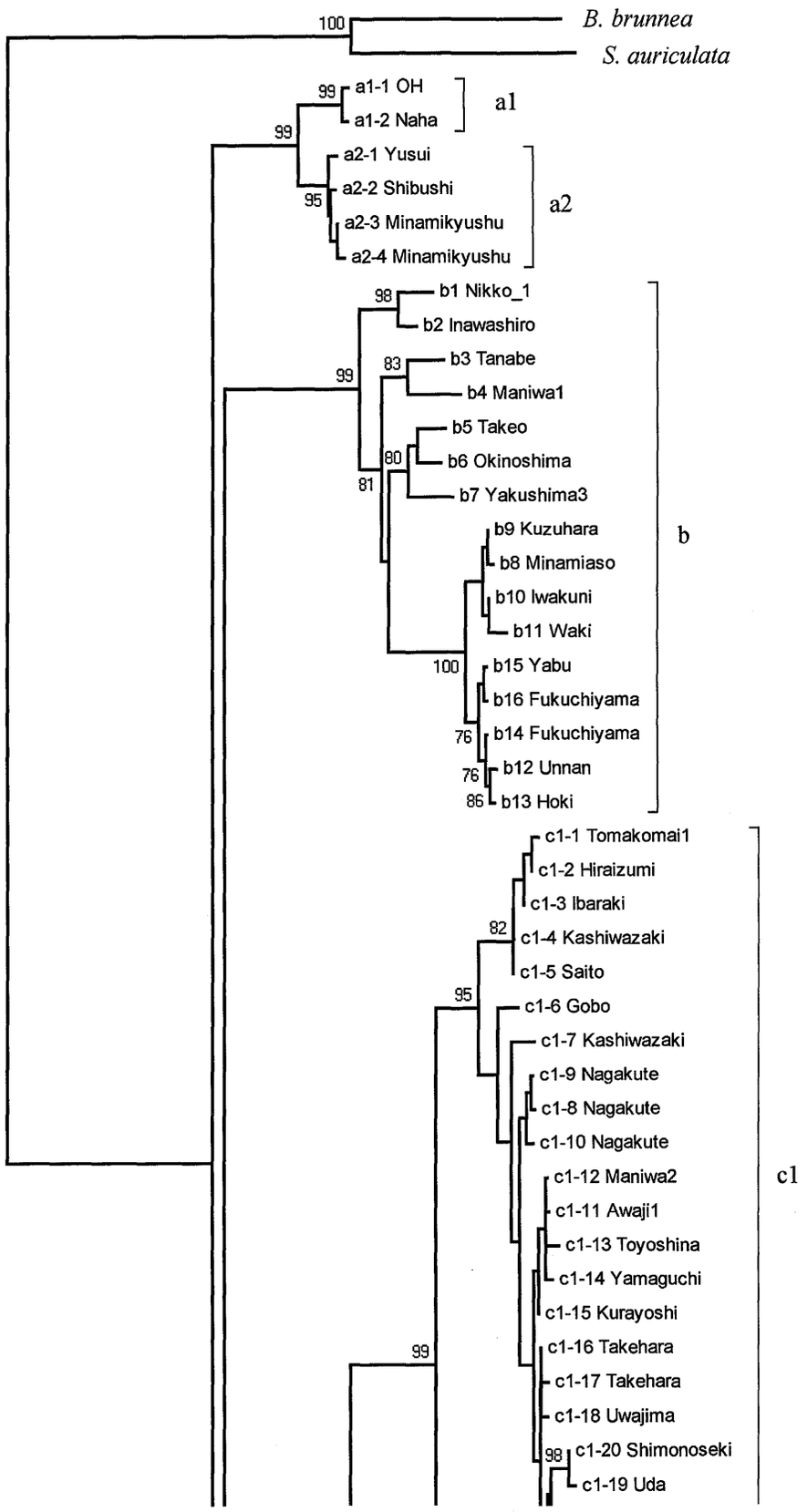


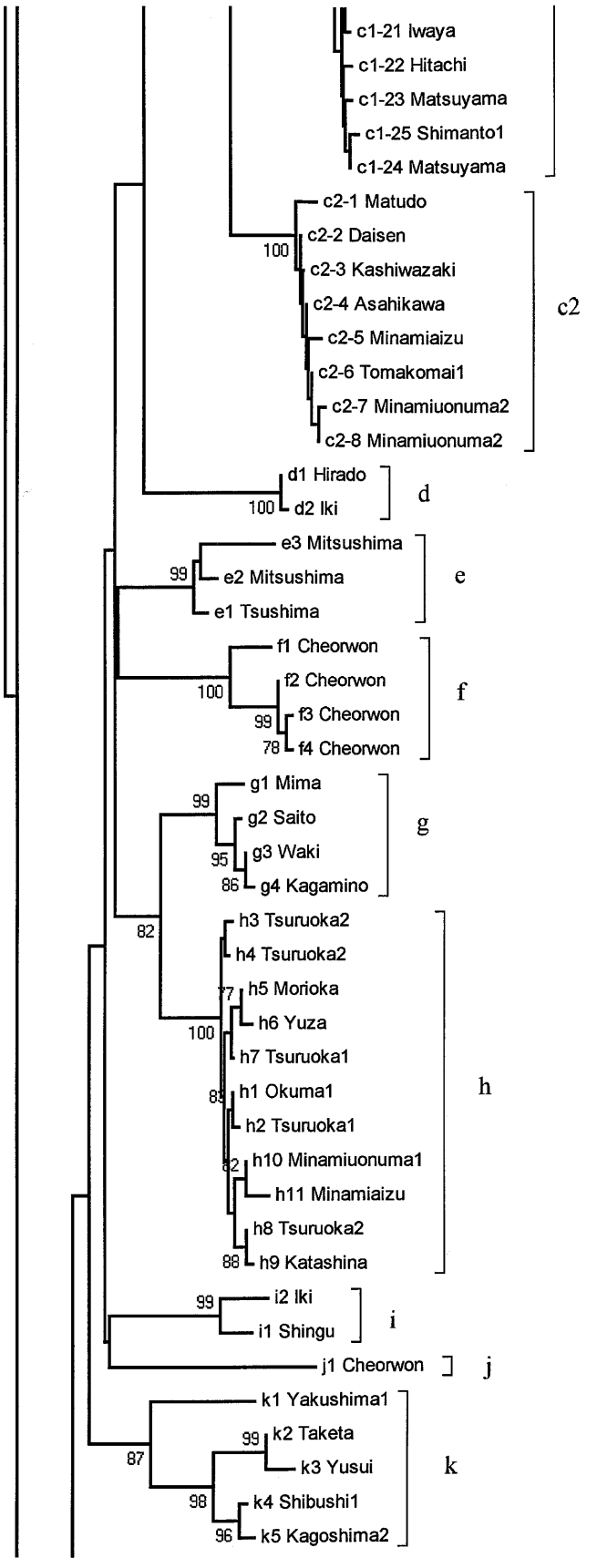
Fig. 11(C). Geographic distributions of genetic groups J-III, -V and –VI.

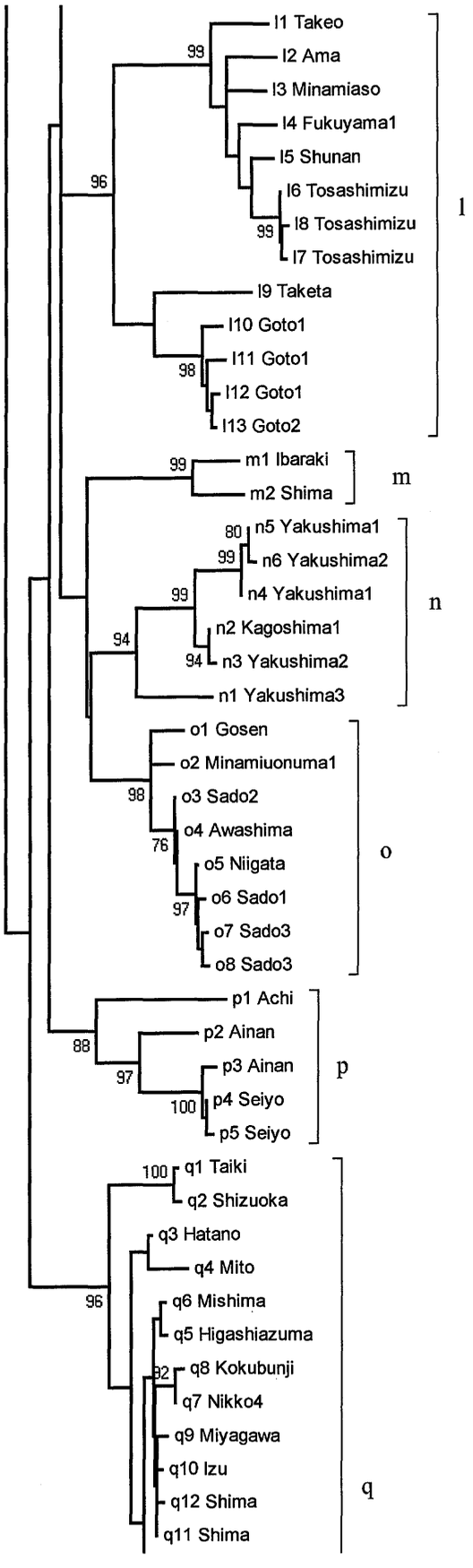

B. brunnea
S. auriculata
S. mediterranes type I
OH
Naha
Yusui
D. ryukyuensis
Unnan
Goto1
Mima
Yuza
Shingu
Tomakomai
Sendai
Awashima
D. japonica
1.0/100
1.0/98
1.0/100
1.0/100
S. mediterranea type II
A1 OH
A2 Naha
A3 Yusui
A4 Minamikyushu
A6 Kagoshima1_2
A5 Kagoshima1_1
A7 Shibushi
clade A
D. ryukyuensis
1.0/100
1.0/100
B1 Tanabe
B2 Tokushima
B3 Unnan
B4 Miki
B5 Inawashiro
B6 Maizuru
B7 Asago
clade B
C1 Ama
C2 Minamiaso
C3 Takeo
C4 Fukuyama2
C5 Taketa1_1
C6 Taketa1_2
C7 Goto1_1
C8 Goto1_2
clade C
1.0/80
1.0/100
D1 Achi
D2 Seiyo
D3 Sendai
D4 Mito
D5 Hadano
D6 Minamiuonuma
D7 Taiki
D8 Mishima
D9 Maibara
D10 Nikko4
clade D
1.0/100
E3 Azumino
E2 Towada
E1 Tomakomai1
E4 Takehara
E5 Shingu
E6 Shimonoseki
E7 Tsushima1
E8 Minamiaizu
E9 Kurayoshi
E10 Cheorwon
E11 Yuza
E12 Tsuruoka2
E13 Sanjo
D. japonica
1.0/99

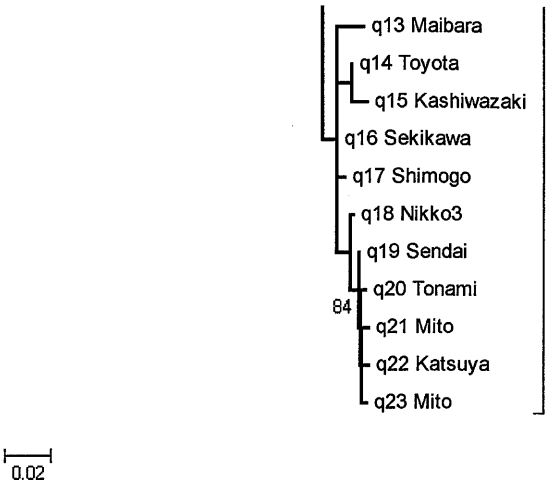


Supplementary figure 1. Detail tree constructed from a Bayesian and Maximum likelihood analysis of 1690 bp from 18S rRNA type I and type II gene.


100
B. brunnea
S. auriculata
99
a1-1 OH
a1-2 Naha
a1
99
a2-1 Yusui
a2-2 Shibushi
95
a2-3 Minamikyushu
a2-4 Minamikyushu
a2
98
b1 Nikko_1
b2 Inawashiro
83
b3 Tanabe
99
b4 Maniwa1
b5 Takeo
80
b6 Okinoshima
81
b7 Yakushima3
b9 Kuzuhara
b8 Minamiaso
b10 Iwakuni
b11 Waki
b
100
b15 Yabu
b16 Fukuchiyama
76
b14 Fukuchiyama
76
b12 Unnan
86
b13 Hoki
c1-1 Tomakomai1
c1-2 Hiraizumi
c1-3 Ibaraki
82
c1-4 Kashiwazaki
c1-5 Saito
95
c1-6 Gobo
c1-7 Kashiwazaki
c1-9 Nagakute
c1-8 Nagakute
c1-10 Nagakute
c1-12 Maniwa2
c1-11 Awaji1
c1-13 Toyoshina
c1-14 Yamaguchi
c1-15 Kurayoshi
99
c1-16 Takehara
c1-17 Takehara
c1-18 Uwajima
98
c1-20 Shimonoseki
c1-19 Uda
c1


c1-21 Iwaya
c1-22 Hitachi
c1-23 Matsuyama
c1-25 Shimanto1
c1-24 Matsuyama
c2-1 Matudo
c2-2 Daisen
c2-3 Kashiwazaki
c2-4 Asahikawa
c2-5 Minamiaizu
c2-6 Tomakomai1
c2-7 Minamiuonuma2
c2-8 Minamiuonuma2
c2
100
d1 Hirado
d2 Iki
100
d
e3 Mitsushima
e2 Mitsushima
e1 Tsushima
99
e
f1 Cheorwon
f2 Cheorwon
f3 Cheorwon
f4 Cheorwon
100
99
78
f
g1 Mima
g2 Saito
g3 Waki
g4 Kagamino
99
95
86
g
82
h3 Tsuruoka2
h4 Tsuruoka2
h5 Morioka
h6 Yuza
h7 Tsuruoka1
h1 Okuma1
h2 Tsuruoka1
h10 Minamiuonuma1
h11 Minamiaizu
h8 Tsuruoka2
h9 Katashina
77
100
85
82
88
h
99
i2 Iki
i1 Shingu
i
j1 Cheorwon
j
k1 Yakushima1
k2 Taketa
k3 Yusui
k4 Shibushi1
k5 Kagoshima2
99
87
98
96
k


l1 Takeo
l2 Ama
l3 Minamiaso
l4 Fukuyama1
l5 Shunan
l6 Tosashimizu
l8 Tosashimizu
l7 Tosashimizu
l9 Taketa
l10 Goto1
l11 Goto1
l12 Goto1
l13 Goto2
l
m1 Ibaraki
m2 Shima
m
n5 Yakushima1
n6 Yakushima2
n4 Yakushima1
n2 Kagoshima1
n3 Yakushima2
n1 Yakushima3
n
o1 Gosen
o2 Minamiuonuma1
o3 Sado2
o4 Awashima
o5 Niigata
o6 Sado1
o7 Sado3
o8 Sado3
o
p1 Achi
p2 Ainan
p3 Ainan
p4 Seiyo
p5 Seiyo
p
q1 Taiki
q2 Shizuoka
q3 Hatano
q4 Mito
q6 Mishima
q5 Higashiazuma
q8 Kokubunji
q7 Nikko4
q9 Miyagawa
q10 Izu
q12 Shima
q11 Shima
q
99
96
99
98
99
80
99
99
94
94
98
76
97
88
97
100
100
96
92



Supplementary figure 2. Detail tree constructed from a Neighbor-joinig analysis of 381 bp from a partial *coI* gene.

Table caption

Table 1. Long PCR and sequencing primers.

Primers	5' to 3'	Primers	5' to 3'
Dugesia japonica		*Dugesia ryukyuensis*	
Long PCR primers			
LA_f_01	ACTGTTGGTGGTTTAACTGGTTTAG	cytb_f_01	GTTATGTTTTGCCTTGAGGTAATATG
LA_rev_01	AAACCTTAATACCAGTAGGAACAGC	LA_rev_07	ATAGCAAACAACATACCCAAGTGAC
LA_f_02	AGTGTTGTTTTGAATTGATTATTGGG	LA_f_07	CTGCTACTATGATTATTGCTGTTCC
		rrnL_rev_01	TACATCGGTCTAAACTCAAATCACG
		rrnL_f_02	CTGTACGAAGGTAGCATAATTACTTGTC
		cytb_rev_02	GAAACAAAACGACTCTTCAAAACAG
Sequencing primers			
ShR_f_02	TTGGTTATTTTCTGTTTCCTGG	cytb_f_01	GTTATGTTTTGCCTTGAGGTAATATG
ShR_f_03	CTGGTCTATTTGGTTGCCCTC	cytb_f_02	CTGTTTTGAAGAGTCGTTTTGTTTC
ShR_f_05	GCACCCACTCCTGTTTCTGC	cytb_rev_02	GAAACAAAACGACTCTTCAAAACAG
ShR_rev_05	GCAGAAACAGGAGTGGGTGC	ShL_f14	TTTAGNTCTGAGGCTGCTTTRAG
ShR_f_06	TTCACTCTTCTACTCTTGTTGTTGC	nd4_f_01	GTGGTTCTTTTCCTATWTTTGGTG
ShR_f_07	TAATTCTAGGAGGTTTAACTAGTTTAAG	nd4_f_02	CAGTTTCTCATGGTTTNRTTTCTTC
ShR_f_11	GGGTTTTCCTTTCGGGTTTAG	coI_f_06	TATGGTATAATGACTGCTCATGG
ShR_rev_12	GAACAAACAAGTAACAAAACCGAA	coI_f_07	ATGCCTCGTCGTATTTGTGA
ShR_f_14	CTTTTATTTATTTGTTCTGAAGTTTTG	coI_rev_05	GCACGAGGAAAAGCCATATC
ShR_f_15	GGATTATTCTTCTATTGACCATTCTG	coI_rev_07	ATCACAAATACGACGAGGCA
ShR_f_16	TAGCTTAGGGTTTTGTTCTCTCAC	nd6_f_01	GTTTNTYTAGTTGYCCTTTNCGTTC
ShR_f_17	CATTTGGCTCCTCTAGGTTGTC	nd5_f_01	GCTCCWACTCCWGTTTCTGC
ShR_f_18	GCTGGGGTATTTTCATATGTTG	nd5_f_02	GTCTTTGTTTGTGATAYAATGTTGG
ShR_f_20	CTTGTTTGTTTTTGGGAGAGTATG	coIII_rev_01	CATGTWGGTTTTGARTGTGC
ShR_f_21	ATAGTTTGGTTAAAATGTATTCTTTG	coIII_rev_04	CAAAAAGAGAAGTACAAATAGGCC
ShR_f_22	TTATGGTGATTGTTTAGATGCTTC	atp6_f_01	CKGCTGGTCATRTNTTKWGTCAT
ShR_f_23	TTCTGTTCTATCTCCTTCTGGG	atp6_rev_01	TGACWMAANAYATGACCAGCMGA
ShR_f_24	GTTATTTCTGTTTATGAGTGTGGC	nd1_f_01	CGTTCTTCTTTTCCTCGTTTTC
ShR_f_26	GATTATTTATGGTTTTGTTTCTGTTTC	nd1_f_02	CTGARGGTGAGAGGGAGTTRGTTTC
ShR_f_27	CCTTTCTTGTGGTTTGTTCCTA	nd1_rev_01	GAAAACGAGGAAAAGAAGAACG
ShR_f_28	GTTCCTTTTTCTTTGCTTTTGTTG	nd1_rev_03	AAACAGMMATTAAAGGRCTYAACCAA
ShR_f_29	TAGTTAATGACAGGTTCGTGTTAGA	coII_f_01	CTGGTCGACAATGGTATTGAG
ShR_f_31	CTCTTTAGGTGGTTTTCTCGATAC	coII_rev_01	CTCAATACCATTGTCGACCAG
ShR_f_32	CCTGGGGATTATGAGCTTTG	nd3_f_03	AGTCGGGAGGTTATATCTGTTTAC
ShR_f_33	GTGTACGGTTGAACAAAAGACG	nd2_f_03	CTCCTCTGACTCTGATTGCG
ShR_f_34	CCTTTTTGTATCACGGGTTTTAG	nd2_f_04	TGGTGTCAATACTTTAGTTATGTTGG
ShR_f_38	CTTATAAGCTGGCATATTTTTTTTCG	ShR_f_49	TTCTGATGTTGTTGGTTTAGACA
ShR_f_40	CAAGTTAGTTAAAAAACAGGCAAATATG	ShR_rev_44	AAAGATTTAACTTGCCGGAAC
ShR_f_41	CAGGCAAATATGAAGGGGG	ShR_rev_47	GAAAACACCTGAAGAACATGAATC
ShR_f_46	ATACCTGTAATACATACTATGATTG	rrnS_rev_01	GAGATTGACGGGCAGTGTG
ShR_rev_46	TAGTATGTATTACAGGTATAATATTATG	rrnL_f_01	CGTGATTTGAGTTTAGACCGATGTA
ShL_f_01	TTGTCTTTTTGTTTAACTATTCGTAA	rrnL_f_02	CTGTACGAAGGTAGCATAATTACTTGTC
ShL_f_02	CTTCTATAGGTTCTTTTGGTTTTAGTAG	rrnL_rev_02	GACAAGTAATTATGCTACCTTCGTACAG
ShL_f_03	CAGTTTCTCATGGTTTTGTTTCT		
ShL_rev_02	CTACTAAAACCAAAAGAACCTATAGAAG		
ShL_rev_03	AGAAACAAAACCATGAGAAACTG		
ShL_rev_04	TCTATCTAACTGACGAACAGAACAA		
ShL_rev_05_2	CTAGTCGCTCAGGGTTATAACCTC		
ShL_rev_06_2	CAATATCCGTAACACAAAAACTAGG		
ShL_rev_07_2	CATACCCAAACCAACATAAGGAC		
ShL_rev_08_2	CTCTTAAAACAATACACCTCCTTGG		
ShL_rev_09	CTACCAACAAGTCAAACACCAAG		
ShL_rev_10	GTAACTACCATAATAAAGACCACGAGA		
ShL_rev_11	GCTTTAGCCTTACTCCCTCC		
ShL_rev_12	TACTCTTACTATGGGAAATCAATTACC		

Table 2. Annotation of the mitochondrial genomes of *D. japonica* and *D. ryukyuensis*. *nd1–6*, NADH dehydrogenase subunits 1–6; *coI–III*, cytochrome *c* oxidase subunits I–III; *atp6*, ATP subunit 6; *cytb*, cytochrome *b*.

Gene	From	To	Size (bp)	Start codon	Stop codon	Gene	From	To	Size (bp)	Start codon	Stop codon
Dugesia japonica						*Dugesia ryukyuensis*					
cytb	1	1158	1158	ATT	TAG	*cytb*	889	1998	1100	TTG	TAG
nd4	1381	2718	1338	ATG	TAA	*nd4*	2221	3579	1359	GTG	TAG
coI	3049	4572	1524	TTG	TAG	*coI*	3835	5382	1548	TTG	TAG
nd6	4649	5092	444	ATA	TAG	*nd6*	5472	5921	450	ATT	TAG
nd5	5096	6679	1584	ATT	TAG	*nd5*	5922	7277	1356	GTG	TAA
trnS2	6700	6770	71			*trnS2*	7630	7688	59		
trnD	6771	6836	66			*trnD*	7692	7744	53		
trnR	6837	6888	52			*trnR*	7751	7823	73		
coIII	6907	7722	816	TTG	TAG	*coIII*	7834	8679	846	TTG	TAA
trnI	7724	7785	62			*trnI*	8692	8745	54		
trnQ	7794	7848	55			*trnQ*	8756	8813	58		
trnK	7849	7909	61			*trnK*	8814	8870	57		
atp6	7912	8547	636	TTG	TAG	*atp6*	8878	9534	642	TTG	TAA
trnV	8549	8606	58			*trnV*	9517	9572	56		
nd1	8593	9495	903	ATT	TAG	*nd1*	9553	10461	909	ATT	TAG
trnW	9500	9565	66			*trnW*	10467	10539	73		
coII	9587	10270	684	ATG	TAA	*coII*	10555	11245	691	ATG	T
trnP	10256	10318	63			*trnP*	12115	12174	60		
nd3	10319	10663	345	TTG	TAG	*nd3*	12178	12519	342	TTG	TAG
trnL2	10674	10743	70			*trnL2*	12539	12626	88		
nd2	10764	11660	897	ATG	TAG	*nd2*	12611	13510	900	ATA	TAG
trnE	11799	11868	70			*trnN*	13995	14070	76		
trnN	11876	11946	71			*trnM*	14475	14534	60		
trnM	12081	12140	60			*trnH*	14679	14743	65		
trnH	12172	12239	68			*trnF*	14763	14824	62		
trnF	12273	12334	62			*rrnS*	14825	15598	774		
rrnS	12335	13109	775			*trnY*	15599	15659	61		
trnY	13110	13169	60			*trnG*	15661	15723	63		
trnG	13170	13229	60			*trnS1*	15724	15790	67		
trnS1	13234	13303	70			*rrnL*	15792	16679	888		
rrnL	13306	14199	894			*trnL1*	16681	16742	62		
trnL1	14200	14260	61			*trnT*	16744	16803	60		
trnT	14261	14317	57								
trnA	14351	14426	76								
trnC	15812	15872	61								

Table 3. Nucleotide divergences of protein coding and ribosomal RNA genes within and between *Dugesia* species. 1; *D.japonica* from Gosen, 2; Unnan, 3; *D.ryukyuensis* from Naha and 4; Shibushi.

nd1	1	2	3
1			
2	0.27		
3	0.31	0.30	
4	0.33	0.31	0.06

nd2	1	2	3
1			
2	0.59		
3	0.61	0.70	
4	0.62	0.68	0.10

nd3	1	2	3
1			
2	0.50		
3	0.47	0.50	
4	0.54	0.55	0.09

nd4	1	2	3
1			
2	0.33		
3	0.46	0.45	
4	0.44	0.45	0.07

nd4l	1	2	3
1			
2	0.30		
3	0.50	0.51	
4	0.45	0.47	0.08

nd5	1	2	3
1			
2	0.44		
3	0.60	0.63	
4	0.61	0.64	0.08

nd6	1	2	3
1			
2	0.78		
3	0.77	0.87	
4	0.91	0.95	0.11

Average	1	2	3
1			
2	0.42		
3	0.51	0.54	
4	0.53	0.56	0.08

cytb	1	2	3
1			
2	0.29		
3	0.29	0.32	
4	0.33	0.33	0.08

coI	1	2	3
1			
2	0.26		
3	0.25	0.28	
4	0.25	0.27	0.07

coII	1	2	3
1			
2	0.59		
3	0.81	0.63	
4	0.89	0.67	0.09

coIII	1	2	3
1			
2	0.42		
3	0.54	0.57	
4	0.52	0.58	0.08

atp6	1	2	3
1			
2	0.46		
3	0.76	0.86	
4	0.76	0.91	0.09

srrn	1	2	3
1			
2	0.39		
3	0.45	0.56	
4	0.42	0.56	0.04

lrrn	1	2	3
1			
2	0.28		
3	0.39	0.41	
4	0.41	0.43	0.04

Table 4. Collection sites, sample size (N) and observed 18S rRNA type II haplotypes and *coI* mitotypes of *Dugesia* species.

Collection site	N	Type II haplotype	*coI* mitotype
1 Medanbetsu, Hokkaido, Pref.	1	E1	c2
2 Asahikawa, Hokkaido, Pref.	1	E1	c2-4
3 Sapporo, Hokkaido, Pref.	1	E1	c1-3
	3		c1-3
4 Tomakomai 1, Hokkaido, Pref.	1	E1	c2-6
	3		c1-1(1), c2-6(2)
5 Tomakomai 2, Hokkaido, Pref.	1	E16	c1-3
	3		c1-3
6 Hirakawa, Aomori, Pref.	1	E1	c2-6
	3		c2-6
7 Towada, Aomori, Pref.	1	E2	
8 Daisen, Akita, Pref.	1	E1	c2-2
	3		c2-2
9 Yuza, Yamagata, Pref.	1	E11	h6
	3		h6
10 Murayama, Yamagata, Pref.	1	D10	q19
	3		q19
11 Tsuruoka 1, Yamagata, Pref.	1	E12	h2
	2		h2(1), h7(1)
12 Tsuruoka 2, Yamagata, Pref.	3	E12	h8(1), h3(1), h4(1)
	1	G6	o4
13 Morioka, Iwate, Pref.	1	E12	h5
	3		h5
14 Hiraizumi, Iwate, Pref.	1	E1	c1-2
	3		c1-2
15 Ichinoseki, Iwate, Pref.	1	D3	q20
	3		q20
16 Rikuzentakata, Iwate, Pref.	1	E1	c2-4
	3		c2-4
17 Sendai, Miyagi, Pref.	1	D3	q19
	3		q19
18 Inawashiro, Fukushima, Pref.	4	B5	b2
19 Okuma, Fukushima, Pref.	1	E12	h1
	3		h1
20 Tamura, Fukushima, Pref.	1	E1	c2-7
	3		c2-7
21 Minamiaizu, Fukushima, Pref.	1	E8	c2-5
	2	E12	c2-5(1), h11(1)
22 Shimogo, Fukushima, Pref.	1	D3	q17
	3		q17
23 Nikko 1, Tochigi, Pref.	1	B5	b1
	1		b1
24 Nikko 2, Tochigi, Pref.	1	D3	q19
	1		q19
25 Nikko 3, Tochigi, Pref.	1	D3	q18
	1		q7
26 Nikko 4, Tochigi, Pref.	1	D10	q7
	1		q7
27 Katashina, Gunma, Pref.	1	D10	q7
	1	E12	h9
28 Higashiazuma, Gunma, Pref.	1	D10	q7
	3		q7

Collection site	N	Type II haplotype	*coI* mitotype
29 Mito, Ibaraki, Pref.	1	D4	q4
	3		q21(2), q23(1)
30 Hitachi, Ibaraki, Pref.	1	E15	c1-22
31 Yorii, Saitama, Pref.	1	D8	q7
	3		q7
32 Matsudo, Chiba, Pref.	1	E1	c2-1
	3		c2-1
33 Kokubunji, Tokyo, Pref.	1	D10	q8
	3		q8
34 Hadano, Kanagawa, Pref.	1	D5	q3
	3		q3
35 Awashimaura, Niigata, Pref.	1	G6	o4
	3		o4
36 Sekikawa, Niigata, Pref.	1	D10	q19
37 Sado 1, Niigata, Pref.	1	G6	o3
	3		o3
38 Sado 2, Niigata, Pref.	1	G6	o3
	3		o3
39 Sado 3, Niigata, Pref.	1	G6	o7
	1		o8
40 Sado 4, Niigata, Pref.	1	G6	o6
	3		o6
41 Sado 5, Niigata, Pref.	1	G6	o3
	2		o3(1), o6(1)
42 Niigata, Niigata, Pref.	3	G6	o5(2), o7(1)
43 Gosen, Niigata, Pref.	3	G5	o1
	1	D3	q14
44 Aga, Niigata, Pref.	1	D10	q17
	3		q17
	1		q16
45 Sanjo, Niigata, Pref.	1	E13	
46 Uonuma, Niigata, Pref.	1	D3	q14
	3		q14(2), q15(1)
47 Minamiuonuma 1, Niigata, Pref.	1	E13	h10
	1	G5	o2
	2	D6	
48 Minamiuonuma 2, Niigata, Pref.	1	E13	c2-7
	3		c2-8
49 Tsunan, Niigata, Pref.	1	E15	q15
50 Kashiwazaki, Niigata, Pref.	1	D10	q15
	2	E1	c1-7
	1	E5	c2-3
51 Itoigawa, Niigata, Pref.	1	G6	o4
	3	D3	q14
52 Azumino, Nagano, Pref.	2	E3	c1-13
53 Iida, Nagano, Pref.	1	D10	q23
	3		q19(2), q23(1)
54 Achi, Nagano, Pref.	2	D1	p1
	2	D10	q15
55 Hokuto, Yamanashi, Pref.	1	E15	c2-4
	3		c2-4
56 Nyuzen, Toyama, Pref.	1	D10	q14
57 Tonami, Toyama, Pref.	1	D10	q20

Collection site	N	Type II haplotype	*coI* mitotype
58 Kanazawa, Ishikawa, Pref.	1	E1	c1-12
	3		c1-12
59 Katsuyama, Fukui, Pref.	1	D3	q19
	3		q19(2), q22(1)
60 Mishima, Shizuoka, Pref.	1	D8	q6
	3		q5(1), q6(2)
61 Izu, Shizuoka, Pref.	1	D3	q10
	2		q10
62 Shizuoka, Shizuoka, Pref.	1	D10	q2
	2		q2
63 Toyota, Aichi, Pref.	1	D3	q14
64 Nagakute, Aichi, Pref.	1	D3	q19
	1	E14	c1-8
	2	E15	c1-9(1), c1-10(2)
65 Toyokawa, Aichi, Pref.	3	E23	g3
66 Gujo, Gifu, Pref.	1	D1	q13
67 Yoro, Gifu, Pref.	1	D9	q13
	3		q13
68 Maibara, Shiga, Pref.	1	D9	q13
	3		q13
69 Miyagawa, Mie, Pref.	1	D10	q7
	3		q7
70 Taiki, Mie, Pref.	1	D7	q1
	3		q1
71 Shima, Mie, Pref.	2	D10	q10
	2	G1	m2
72 Shingu, Wakayama, Pref.	1	E25	i1
	1	E26	i1
	2	E5	c1-3
73 Tanabe, Wakayama, Pref.	1	B1	b3
	3		b3
74 Gobo, Wakayama, Pref.	3	E1	c1-6(1), c1-21(2)
	1	E27	i1
75 Uda, Nara, Pref.	1	E1	c1-19
	3		c1-19
76 Maizuru, Kyoto, Pref.	1	D10	q18
	2	B5	b14
	1	B6	b14
77 Fukuchiyama, Kyoto, Pref.	1	E21	g3
	3	B5	b14(2), b16(1)
78 Kameoka, Kyoto, Pref.	1	D3	q19
	3		q19
79 Ibaraki, Osaka, Pref.	3	E17	c1-3
	1	G2	m1
80 Tanba, Hyogo, Pref.	1	E4	c1-3
	1		c1-3
81 Asago, Hyogo, Pref.	1	B7	b16
	1	B5	b16
	1	E24	i1
82 Yabu, Hyogo, Pref.	3	B5	b15
	1	E1	c1-20
83 Awaji1, Hyogo, Pref.	4	E1	c1-11

Collection site	N	Type II haplotype	*coI* mitotype
84 Awaji2, Hyogo, Pref.	1	E4	c1-21
	3		c1-12
85 Waki, Okayama, Pref.	1	E21	g4
	2	B5	b11
86 Kagamino, Okayama, Pref.	1	E21	g4
	3		g4
87 Maniwa 1, Okayama, Pref.	1	B5	b4
	1		b4
88 Maniwa 2, Okayama, Pref.	1	E4	c1-12
	1		c1-12
89 Fukuyama 1, Hiroshima, Pref.	1	C4	l4
90 Fukuyama 2, Hiroshima, Pref.	1	B5	b4
	1	E21	g3
91 Shobara, Hiroshima, Pref.	1	B5	b4
	1	E22	g3
92 Takehara, Hiroshima, Pref.	3	E4	c1-16(1), c1-17(1), c1-27(1)
	1	B5	b8
93 Hiroshima, Hiroshima, Pref.	2	B5	b9
	2	E21	g3
94 Chizu, Tottori, Pref.	1		g3(1)
95 Chizu2, Tottori, Pref.	1	E21	c1-12
	3		c1-12(2), g3(1)
96 Kurayoshi, Tottori, Pref.	1	E1	c1-14
	2		c1-15
97 Hoki, Tottori, Pref.	3	B5	b13(1), b14(2)
98 Okinoshima, Shimane, Pref.	1	B5	b6
	3		b6
99 Ama, Shimane, Pref.	1	C1	l2
	2	C4	l2
100 Unnan, Shimane, Pref.	1	B3	b12
	3		b12
101 Hamada, Shimane, Pref.	1	B3	b12
	2	B5	b12
	3	E27	i1
102 Yoshika, Shimane, Pref.	1	E1	
103 Iwakuni, Yamaguchi, Pref.	1	B5	b10
	1	E20	g3
	2	E21	g3
104 Shunan			l5
105 Yamaguchi, Yamaguchi, Pref.	1	E1	c1-14
	3		c1-14
106 Mine, Yamaguchi, Pref.	2		c1-12
107 Shimonoseki, Yamaguchi, Pref.	1	E6	c1-20
	3		c1-20
108 Miki, Kagawa, Pref.	1	B4	b4
109 Tokushima, Tokushima, Pref.	1	B2	b
110 Mima, Tokushima, Pref.	2	E21	g1
	2	B5	b4
111 Minami, Tokushima, Pref.	1	B4	b4
	1		b4
112 Saijo, Ehime, Pref.	1	B5	b4
	3		b4(1), b9(2)

Collection site	N	Type II haplotype	*coI* mitotype
113 Imabari, Ehime, Pref.	1	E28	i1
	3		i1
114 Matsuyama, Ehime, Pref.	1	E4	c1-24
	3		c1-23(1), c1-24(2)
115 Uwajima, Ehime, Pref.	1	E4	c1-18
	3		c1-18
116 Seiyo, Ehime, Pref.	1	D2	p4
	2		p4(1), p5(1)
117 Ainan, Ehime, Pref.	1	D2	p2
	2		p3
118 Shimanto, Kochi, Pref.	3	E1	c1-25
	1	E27	i1
119 Tosashimizu, Kochi, Pref.	1	C4	l6
	3		l6(1), l7(1), l8(1)
120 Kurate, Fukuoka, Pref.	2	B5	b8
	1		c1-3(1)
121 Takeo, Saga, Pref.	3	B5	b5
	1	C2	l1
122 Iki, Nagasaki, Pref.	3	E6	d1(1), d2(2)
	1	E29	i2
123 Tsushima1, Nagasaki, Pref.	1	E7	e1
	3		e1(2), e3(1)
124 Tsushima2, Nagasaki, Pref.	1	E7	e2
	3		e3
125 Hirado, Nagasaki, Pref.	1	E6	d1
	3		d1
126 Goto 1, Nagasaki, Pref.	2	C7	l10(1), l12(1)
	1	C8	l11
127 Goto 2, Nagasaki, Pref.	1	C7	l10
	2	C8	l10(1), l13(1)
128 Taketa, Oita, Pref.	1	C5	l9
	1	C6	l9
	1	E19	g3
	1	F1	k2
129 Minamiao, Kumamoto, Pref.	1	C2	l3
	1	B5	b8
	2	F7	k2
130 Saito, Miyazaki, Pref.	2	E19	g2(1), g3(1)
	2		c1-5
131 Yusui, Kagoshima, Pref.	3	A3	a2-1
	1	F3	k3
132 Shibushi 1, Kagoshima, Pref.	1	A7	a2-2
	3		a2-2
133 Shibushi 2, Kagoshima, Pref.	3	A7	a2-2
	1	F2	k4
134 Kagoshima 1, Kagoshima, Pref.	3	A5	a2-3
	3	A6	n2
	3		a2-3
135 Kagoshima 2, Kagoshima, Pref.	3	F6	k5
136 Minamikyushu, Kagoshima, Pref.	4	A3	a2-3
137 Yakushima 1, Kagoshima, Pref.	5	F5	k1
	5	G3	n5

	Collection site	N	Type II haplotype	*coI* mitotype
138	Yakushima 2, Kagoshima, Pref.	1	F5	k5
		6	G3	n2(1), n3(3), n5(1), n6(1)
139	Yakushima 3, Kagoshima, Pref.	8	B5	b7
		1	G3	n1
140	Naha, Okinawa, Pref.	2	A2	a1-2
141	Cheorwon, Koria	2	E10	f1(1), f2(1)
		1	E29	j1
		4		f1(7), f3(1), f4(1), j1(1)
	OH strain	2	A1	a1-1

Table 5. PCR and sequencing primers for 18S rRNA type I and type II, and *coI*.

Primers	Sequence (5' to 3')	Reference
PCR primers		
18S rRNA type I of *D. japonica* and *D. ryukyuensis*		
1F	TACCTGGTTGATCCTGCCAGTAG	(Salvador et al., 1996)
typeI_R	AAGGAGATTTGAATTACCTTCGCG	(in this study)
typeI_F	CGCGAAGGTAATTCAAATCTCCTT	(in this study)
9R	GATCCTTCCGCAGGTTCACCTAC	(Salvador et al., 1996)
18S rRNA type II of *D. japonica*		
1F		
typeII_3R	AGTAGATTGTCTGCTGATGGTCTC	(in this study)
typeII_2F	TCCTTCTCCGTCGTGTATATTG	(in this study)
9R		
18S rRNA type II of *D. ryukyuensis*		
1F		
type_II_4R	GACAGTGTTTRACAAGATTACCCAAC	(in this study)
type_II_4F	GAGACTCTGACTTGCTAAATAGTGG	(in this study)
9R		
coI		
CO1_f	AGCTGCAGTTTTGGTTTTTTGGACATCCTGAGGT	(Bessho et al., 1992a.b)
CO1_r	ATGAGCAACAACATAATAAGTATCATG	(Bessho et al., 1992a.b)
Sequencing primers		
18S rRNA type I		
18S_3F	GTTCGATTCCGGAGAGGGA	(Salvador et al., 1996)
typeI_R	AAGGAGATTTGAATTACCTTCGCG	(in this study)
18S_7F	GCAATAACAGGTCTGTGATGCCC	(Salvador et al., 1996)
18S_7R	GGGCATCACAGACCTGTTATTGC	(Salvador et al., 1996)
18S_9R		
18S rRNA type II		
18S_3F		
typeII_2R	ATATACACGACGGAGAAGGAAGA	(in this study)
18S_10F	GATAGCTCTTTCTTGATTCGGTG	(in this study)
18S_7R		
coI		
CO1_f		
CO1_r		

Table 6. Intraspecific and interspecific nucleotide sequence divergence of 18S rRNA type I and type II in *Dugesia* species.

		1	2	3	4	5	6	7	8	9
1	typeI of *D. ryukyuensis*	0								
2	typeI of *D. japonica*	0.002-0.004	0-0.003							
3	clade A			0-0.007						
4	clade B			0.037-0.045	0-0.009					
5	clade C			0.038-0.049	0.021-0.027	0-0.011				
6	clade D			0.049-0.054	0.027-0.038	0.020-0.030	0-0.012			
7	clade E			0.044-0.054	0.024-0.035	0.014-0.023	0.015-0.026	0-0.010		
8	clade F			0.044-0.051	0.028-0.034	0.016-0.025	0.016-0.025	0.008-0.015	0-0.004	
9	clade G			0.044-0.052	0.028-0.037	0.015-0.026	0.017-0.028	0.012-0.024	0.009-0.019	0-0.008

Table 7. Intraspecific and interspecific nucleotide sequence divergence of *coI* in *Dugesia* speices.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	a-1	0																		
2	a-2	0.03-0.05	0-0.01																	
3	b	0.17-0.20	0.16-0.21	0-0.11																
4	c-1	0.18-0.24	0.18-0.24	0.21-0.34	0-0.08															
5	c-2	0.17-0.20	0.18-0.20	0.24-0.30	0.06-0.12	0-0.02														
6	d	0.19-0.21	0.18-0.20	0.23-0.28	0.15-0.18	0.13-0.16	0													
7	e	0.18-0.21	0.17-0.21	0.18-0.24	0.12-0.21	0.13-0.18	0.13-0.18	0.02-0.05												
8	f	0.20-0.22	0.19-0.20	0.20-0.27	0.15-0.21	0.14-0.16	0.14-0.16	0.09-0.17	0.01-0.05											
9	g	0.16-0.17	0.16-0.18	0.20-0.24	0.13-0.18	0.12-0.16	0.15-0.17	0.09-0.14	0.13-0.16	0-0.03										
10	h	0.15-0.18	0.15-0.19	0.16-0.23	0.15-0.21	0.13-0.17	0.15-0.18	0.10-0.15	0.12-0.16	0.07-0.10	0-0.03									
11	i	0.18-0.19	0.17-0.19	0.18-0.27	0.15-0.19	0.14-0.17	0.15-0.18	0.10-0.15	0.16-0.19	0.16-0.19	0.14-0.17	0.04								
12	j	0.19-0.19	0.19-0.19	0.19-0.22	0.18-0.24	0.18-0.20	0.25-0.26	0.17-0.21	0.21-0.22	0.18-0.19	0.15-0.17	0.17-0.17	0							
13	k	0.18-0.21	0.19-0.23	0.20-0.26	0.18-0.24	0.17-0.23	0.14-0.19	0.13-0.17	0.17-0.22	0.15-0.19	0.11-0.17	0.15-0.18	0.19-0.23	0.01-0.11						
14	l	0.19-0.22	0.19-0.21	0.19-0.27	0.17-0.25	0.19-0.24	0.19-0.22	0.12-0.22	0.14-0.24	0.14-0.22	0.14-0.21	0.17-0.23	0.21-0.24	0.15-0.22	0-0.16					
15	m	0.21-0.22	0.19-0.20	0.19-0.25	0.19-0.23	0.19-0.22	0.17-0.19	0.14-0.18	0.19-0.23	0.17-0.19	0.17-0.20	0.13-0.16	0.22-0.23	0.14-0.18	0.15-0.21	0.05				
16	n	0.17-0.21	0.16-0.21	0.16-0.23	0.19-0.24	0.18-0.24	0.16-0.19	0.12-0.15	0.15-0.22	0.16-0.21	0.12-0.17	0.16-0.19	0.20-0.22	0.14-0.17	0.14-0.21	0.11-0.18	0-0.10			
17	o	0.16-0.20	0.14-0.17	0.16-0.25	0.15-0.20	0.17-0.21	0.16-0.18	0.12-0.17	0.17-0.21	0.13-0.18	0.12-0.18	0.14-0.17	0.17-0.20	0.14-0.20	0.11-0.19	0.10-0.12	0.09-0.13	0-0.05		
18	p	0.16-0.19	0.17-0.19	0.18-0.25	0.19-0.25	0.19-0.23	0.16-0.19	0.15-0.19	0.16-0.20	0.14-0.18	0.12-0.18	0.18-0.22	0.20-0.22	0.13-0.20	0.14-0.23	0.15-0.21	0.12-0.20	0.12-0.18	0-0.12	
19	q	0.13-0.17	0.12-0.16	0.16-0.25	0.17-0.26	0.16-0.22	0.13-0.20	0.13-0.20	0.15-0.19	0.15-0.19	0.15-0.20	0.16-0.20	0.18-0.21	0.15-0.21	0.12-0.22	0.14-0.18	0.15-0.21	0.11-0.16	0.12-0.18	0-0.08

Table 8. Genetic groups on the basis of 18S rRNA type II haplotypes and *coI* mitotypes in Japanese *Dugesia* planarians.

	Type II haplotype	*coI* mitotype	Genetic group
D. ryukyuensis	A	a1,2	R
D. japonica	B	b	J-IV
	C	l	J-III
	D	p, q	J-II
	E	c1,2,d～j	J-I
	F	k	J-V
	G	m, n, o	J-VI

Acknowledgments

I am grateful to are grateful to Drs Y. Kikkawa, S. Hamada, H. Suzuki (Hokkaido Univ), N. Miyashita (Kagawa Univ), Sang-Hoon Han (National Institute of Biological Resources, Korea), T. Uchiuni, S. Hamaguchi and M. Sakaizumi (Niigata Univ.) for their helpful comments on this study.

References

Avise JC (1994) Molecular Markers, Natural History and Evolution. Chapman & Hall, New York

Avise JC (2000) Phylogeography. Harvard Univ. Press.

Baguñà J, Riutort M (2004a) The dawn of bilaterian animals: the case of acoelomorph flatworms. Bioessays 26:1046–1057.

Baguñà J, Riutort M (2004b) Molecular phylogeny of the Platyhelminthes. Can J Zool 82:168–193.

Bessho Y, Ohama T, Osawa S (1992a) Planarian Mitochondria I. Heterogeneity of Cytochrome c Oxidase Subunit I Gene Sequences in the Freshwater Planarian, Dugesia japonica. J Mol Evol 34:324-330.

Bessho Y, Ohama T, Osawa S (1992b) Planarian Mitochondria II. The unique genetic code as deduced from cytochrome C oxidase subunit i gene sequences. J Mol Evol 34:331–335.

Boore JL, Brown WM (1998) Big trees from little genomes: mitochondrial gene order as a phylogenetic tool. Curr Opin Genet Dev 8:668–674.

Ehlers U (1985) Das Phylogenetische System der Plathelminthes Stuttgart. Gustav Fischer, New York.

Ehlers U (1986) Comments on a phylogenetic system of the Platyhelminthes. Hydrobiologia 132:1–12.

Kawakatsu M, Oki I, Tamura S and Sugino H (1976) Studies on the morphology, karyology and taxonomy of the Japanese freshwater planarian *Dugesia japonica* Ichikawa et Kawakatsu, with a description of anew subspecies, *Dugesia japonica ryukyuensis* subspec. nov. Bull . Fuji

Women's College, no 14, ser. II : 81-126.

Kawakatsu M, Oki I, Tamura S (1995) Taxonomy and geographical distribution of *Dugesia japonica* and *D. ryukyuensis* in the Far East. Hydrobiologia 305:55-61

Gissi C, Iannelli F, Pesole G (2008) Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity 101:301–320.

Gutell RR, Cannone JJ, Konings D, Gautheret D (2000a) Predicting U-turns in ribosomal RNA with comparative sequence analysis. J Mol Biol 300:791–803.

Gutell RR, Cannone JJ, Shang Z, Du Y, Serra MJ (2000b) A story: unpaired adenosine bases in ribosomal RNAs. J Mol Biol 304:335–354.

Huyse T, Plaisance L, Webster BL, Mo TA, Bakke TA, Bachmann L, Littlewood DTJ (2007) The mitochondrial genome of Gyrodactylus salaries (Platyhelminthes: Monogenea), a pathogen of Atlantic salmon (Salmo salar). Parasitology 134:739–747.

Huyse T, Buchmann K, Littlewood DT (2008) The mitochondrial genome of Gyrodactylus derjavinoides (Platyhelminthes: Monogenea) a mitogenomic approach for Gyrodactylus species and strain identification. Gene 417:27-34.

Jondelius U, Ruiz-Trillo I, Baguñà J, Riutort M (2002) The Nemertodermatida are basal bilaterians and not members of the Platyhelminthes. Zool Sci 31: 201–215.

Larsson K, Jondelius U (2008) Phylogeny of Catenulida and support for Platyhelminthes. Org Divers Evol 8 (5):378–387.

Lázaro E M., Sluys R, Pala M, Stocchino G A, Baguñà J, Riutort M (2009) Molecular barcoding and phylogeography of sexual and asexual freshwater planarians of the genus *Dugesia* in the Western Mediterranean (Platyhelminthes, Tricladida, Dugesiidae). Mol Phy Evo 52: 835–845

Littlewood DTJ, Cribb TH, Olson PD, Bray RA (2001) Platyhelminth phylogenetics - a key to understanding parasitism? Belg J Zool 131(Suppl. 1):35–46.

Lockyer AE, Olson PD, Littlewood DTJ (2003) Utility of complete large and small subunit rRNA genes in resolving the phylogeny of the Neodermata (Platyhelminthes): implications and a review of the cercomer theory. Biol J Linn Soc 78:155–171.

Marta A ´ lvarez-Presas, Jaume Bagun˜a`, Marta Riutort (2008) Molecular phylogeny of land and freshwater planarians (Tricladida, Platyhelminthes): From freshwater to land and back. Mol Phy Evo 47: 555–568.

Park JK, Kim KH, Kang S, Kim W, Eom KS, Littlewood DT (2007) A common origin of complex life cycles in parasitic flatworms: evidence from the complete mitochondrial genome of *Microcotyle sebastis* (Monogenea: Platyhelminthes). BMC Evol Biol 7:11.

Perkins EM, Donnellan SC, Bertozzi T, Whittington ID (2010) Closing the mitochondrial circle on paraphyly of the Monogenea (Platyhelminthes) infers evolution in the diet of parasitic flatworms. Int J Parasitol 40:1237–1245.

Perna NT, Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol

41(3):353–358.

Philippe H, Brinkmann H, Copley RR, Nakano H, Poustka A J, Wallberg A, Pesterson K J, Telford M J (2011) Acoelomorph flatworms are deuterostomes related to *Xenoturbella*. Nature 470, 255–258.

Ruiz-Trillo I, Riutort M, Littlewood DT, Herniou EA, Baguñà J (1999) Acoel flatworms: earliest extant bilaterian Metazoans, not members of Platyhelminthes. Science 283:1919–1923.

Ruiz-Trillo M, Riutort M, Fourcade HM, Baguna J, Boore JL (2004) Mitochondrial genome data support the basal position of Acoelomorpha and the polyphyly of the Platyhelminthes. Mol Phylogenet Evol 33:321–332.

Salvador C, Gonzalo G, Caries R, Jaume B, Marta R (1996) Evidence that Two Types of 18s rDNA Coexist in the Genome of *Dugesia*. (*Schmidtea*) *mediterranea* (Platyhelminthes, Turbellaria, Tricladida). Mol Biol Evo. 13(6)824-832.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512-526.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24(8):1596–1599.

Tamura S, Yamamoto K, Takai M, Oki I, Kawakatsu M (1998) Karyology and biogeography of *Dugesia japonica* and *Dugesia ryukyuensis* in Kyushu, southern Japan. Hydrobiologia 383:321-327.

Tamura S, Oki I, Kawakatsu M (1991) Karyological and taxonomic studies

of *Dugesia japonica* from the Southwest Islands of Japan-II. Hydrobiologia 227: 157-162.

Tanabe AS (2007) Kakusan: a computer program to automate the selection of a nucleotide substitution model and the configuration of a mixed model on multilocus data. Molecular Ecology Notes 7(6):962-964.

Thompson J.D, Higgins D.G, Gibson T.J (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 2:24673–4680.

Webster BL, Rudolfova J, Horak P, Littlewood DT (2007) The complete mitochondrial genome of the bird schistosome *Trichobilharzia regenti* (Platyhelminthes: Digenea), causative agent of cercarial dermatitis. J Parasitol 93 (3):553-561.

Wolstenholme DR (1992) Animal mitochondrial DNA: structure and evolution. Int Rev Cytol 141:173–216.

Zarowiecki M.Z., Huyse T, Littlewood D.T.J (2008) Making the most of mitochondrial genomes – Markers for phylogeny, molecular ecology and barcodes in *Schistosoma* (Platyhelminthes: Digenea) Gene 417: 27–34