Repetitive sequences in the lamprey mitochondrial DNA control region and speciation of *Lethenteron*

Kazunori Okada[1], Yuji Yamazaki[2], Shinichi Yokobori[3] and Hiroshi Wada[1]

[1] Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba 305-8572 Japan

[2] Department of Biology, Faculty of Science, Toyama University, 3190 Gofuku, Toyama 930-8555, Japan

[3] Department of Molecular Biology, School of Life Science, Tokyo University of Pharmacy and Life Science, 1432 Horinouchi, Hachioji, Tokyo 192-0392, Japan

Author for Correspondence: Hiroshi Wada[1]

Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba 305-8572 JAPAN

Tel & Fax: +81-29-853-4671

E-mail: hwada@biol.tsukuba.ac.jp

Abstract

The sequence of the mitochondrial DNA control region was examined in four species of lamprey in the genus *Lethenteron*. The 3'-half of the control region contains highly variable repeat sequences, showing variation in both copy number and nucleotide sequence, even within local populations. Detailed analyses of the sequences of the repeats allowed us to deduce that slipped-strand mispairing during DNA replication, accompanied by a high rate of substitutions and indels, was primarily responsible for the variation in the repeats. We also found that some cases might be better explained by gene conversion, due to intermolecular recombination. Based on the observed variable nature of the mitochondrial control region, we searched for molecular markers in mitochondrial DNA, because there are few fixed genetic markers for distinguishing between *Lethenteron japonicum* and *L. kessleri*. However, we found no reliable markers in the control region. No fixed substitution was observed in intron sequences of the nuclear gene *SoxD*. Thus, these two species likely diverged quite recently and may possess only a limited number of fixed genetic loci.

Introduction

Sequence variation in mitochondrial DNA (mtDNA) has been widely used for molecular phylogenetic studies. Within mtDNA, the control region has the highest evolutionary rate, and serves as a molecular marker for examining relatively recent events, such as among populations (e.g. Avise et al., 1987; Avise, 1991). The control region shows sequence variation in nucleotide substitutions, indels, and the numbers of tandem repeat sequences. Numerous studies have documented copy number variation within populations, and sometimes in single individuals (*i.e.*, heteroplasmy)(Ludwig et al., 2000; Hoarau et al., 2002; Mjelle et al., 2008).

Most of the copy number variation is thought to arise from slipped-strand mispairing during mtDNA replication (Buroker et al., 1990; Broughton and Dowling, 1994; Mundy and Helbig, 2004). This strand slippage is facilitated by the secondary structure of the repeat sequence. In addition to slipped-strand mispairing, recombination accounts for some of the sequence variation among repeats (Hoarau et al., 2002; Mjelle et al., 2008).

Lamprey mtDNA contains two repeat regions within the control regions, NC1 and NC2 (Lee and Kocher, 1995). Recently, White and Martin (2009) analyzed the copy number variation of the least brook lamprey, *Lampetra aepyptera*, and reported that

the copy number variation in NC1 was due to slipped-strand mispairing.

Besides their phylogenetic position as a basal group of vertebrates, lampreys are an interesting biological group in terms of speciation. All lamprey species breed in fresh water, where they spend several years as suspension or detritus feeders (Hardisty and Potter, 1971). This stage is known as the ammocoetes larval stage. After metamorphosis, while some species parasitize fish and other animals, other species do not feed after metamorphosis, and breed within several months. These non-parasitic species are believed to have evolved several times from ancestral parasitic species (Zanandrea, 1959; Hardisty and Potter, 1971; Hubbs and Potter, 1971; Vladykov and Kott, 1979; Potter, 1980). Variation in the lamprey life history may be one of the factors inducing speciation in lampreys.

The present study focused on four *Lethenteron* species around Japan: *Lethenteron japonicum*, *L. kessleri*, and two cryptic species of *L.* sp. N and *L.* sp. S (Yamazaki and Goto, 1998; Yamazaki et al., 2006). Of these, *L. japonicum* is parasitic species, which grow in the sea after metamorphosis, while *L. kessleri* and the two cryptic species of *L.* sp. N and *L.* sp. S have an entirely freshwater life cycle and non-parasitic. Molecular phylogenetic studies indicated that *L.* sp. S. diverged from the most basal node, and *L.* sp. N., *L. japonicum*, and *L. kessleri* form a monophyletic clade

(Yamazaki et al., 2006). Although several fixed nucleotide polymorphisms in the cytochrome oxidase subunit I gene (*COI*) have been identified in *L.* sp N. and *L.* sp. S., no fixed nucleotide polymorphisms distinguish *L. japonicum* from *L. kessleri* (Yamazaki et al., 2006). To date, only one fixed allele, of malate dehydrogenase 3 (MDH3), has been recognized as a distinct molecular characteristic distinguishing *L. japonicum* and *L. kessleri* (Yamazaki and Goto, 1998).

Because *L. japonicum* is parasitic and *L. kessleri* is non-parasitic, these two species are readily distinguished as adults. However, the ammocoetes larvae of these species are quite difficult to distinguish, based on morphology. Additionally, some individuals of *L. japonicum* have been reported to remain in fresh water (Yamazaki et al., 1998). Thus, genetic markers that distinguish these species are desirable.

In this study, we characterized the molecular structure of the mitochondrial control regions of the Japanese lamprey species, seeking molecular markers to identify *Lethenteron* species. We report the highly variable nature of the NC2 repeats (Lee and Kocher, 1995) in *Lethenteron*. A detailed analysis of the sequence of the repeat arrays allowed us to trace some of the evolutionary history of the generation of variable repeats. However, we could not recover a phylogenetic signal that distinguished *L. japonicum* and *L. kessleri*. These two species likely diverged too recently to detect

genetic fixed genetic loci in mtDNA.

Materials and Methods

Amplification and sequencing of the mtDNA control region and SoxD intron

The sampling localities and numbers of individuals of four species of *Lethenteron* examined are summarized in Table 1. In the present study, we regard *L.* sp. S., as a species of *Lethenteron* following our previsous studies, however, as described in Yamazaki et al. (2006), the classification of the genera *Lethenteron* and *Lampetra* should be re-examined in the future studies. DNA was extracted from body tissues using a DNeasy Kit (QIAGEN). The primers used to amplify the NC2 region of mtDNA control region and *SoxD* are shown in Table 2. These amplified approximately 700 bp of the mtDNA control region and 400 bp of the *SoxD* intron. The nucleotide sequences were determined directly from the PCR fragments after treatment with ExoSAP (GE Healthcare) by using ABI PRISM 377 or 310.

Sequencing the complete mtDNA of L. sp. S.

The complete mtDNA of specimen Ss3 was sequenced. First, the mtDNA was amplified in six fragments, using the six primer pairs (Table 2). Fragments generated were subsequently sequenced by primer-walking.

Sequence analyses

The sequences were compared using the Genetyx software (Genetyx). A network analysis of nucleotide substitutions was performed using TCS 1.21 (Clement et al., 2000). The secondary structure of the DNA was analyzed using MFOLD (Zuker, 2003).

Results and Discussion

Characterization of the mtDNA control region

To examine the genetic structure of *Lethenteron*, we analyzed the sequence variation in NC2 between $tRNA^{Glu}$ and *cytB* (Lee and Kocher, 1995). In this region, repeat sequences have already been reported for *Petromyzon marinus* (Lee and Kocher, 1995) and *Lampetra fluviatilis* (Delarbre et al., 2000), and we found repeat sequences in the four Japanese species (Fig. 1). We sequenced the repeat region from 41 specimens of the four species, and found that the repeats were highly variable in terms of both copy number and the nucleotide sequences of the repeat arrays (Fig. 2A-C). The copy number and sequences were variable even within local populations, in contrast to a report that the NC1 repeat of *Lampetra aepyptera* in the Unites States lacked polymorphisms within local populations (White and Martin, 2009). We did not detect any heteroplasmy, and PCR amplification always resulted in a single DNA band, which was sequenced directly. Detailed analyses of the repeat sequences reveal highly dynamic evolution of the repeat sequences within populations, as described below.

Repeat dynamics within and among species

Before examining the dynamic nature of the repeats, we classified the repeat sequences

based on network analysis of the repeat sequence (Fig. 2A, B). Because the network analysis indicate that A3 and B2 form nodes of the network (Fig. 2B), we classified the repeat sequences into two types based on the 22nd position of the repeat: T in the A type and C in the B type. A3 and B2 were also observed in *L. fluviatilis* (Delarbre et al., 2000) and *P. marinus* (Lee and Kocher, 1995), respectively. The rest of the repeat sequence types were derived via distinct substitutions or indels. All of the repeat sequences possessed multiple T sequences at their 3'-end, and this number was quite variable. We identified the sequences of the repeat arrays in the form type-number of repeats. For example, A3-5 has the A3 type sequence with five Ts at the 3'-end (Fig. 2A). The published sequence of *P. marinus* mtDNA is B2-7, while those of *L. fluviatilis* are A4-6 and A3-5, 8, and 7 (Fig. 2C).

As shown in Figure 2C, the number of repeats is not fixed within species, with six or seven repeats in *L. japonicum*, three to eight repeats in *L.* sp. N., and four or five repeats in *L.* sp. S. Although some individuals within populations possess the same copy number of repeat arrays, the internal architecture of the repeat arrays was quite different. Among the 41 individuals, only two groups shared identical repeat architecture (js5/ji1/ji2 and Nk5/Nk8). Some individuals showed traces of recent events that led to the copy number variation. One example was js4 (*L. japonicum* from

Shiribeshi), in which two arrays of the repeat (B2-6 and A1-5) were duplicated in the third to sixth repeats (Fig. 2C). This pattern of repeats supports the occurrence of slipped-strand mispairing after the two repeat units unfolded, and subsequently two arrays were inserted in the repeats.

Additionally, nk3 shows evidence of slipped-strand mispairing. It has two arrays of the A9 type sequence at the 3'-end of the repeat. Because A9 required two substitutions from the ancestral A3, these were not likely to have emerged simultaneously in the third and fourth repeats, but more likely arose due to slipped-strand mispairing. Similarly, the two arrays of A5 at the 5' end of Ns2 also likely arose via slipped-strand mispairing.

These two cases indicated that slipped-strand mispairing has inserted either one or two copies of the repeat array. Insertion of a single copy is due to folding of a single array of the repeat, while insertion of two copies is due to the folding of the two arrays. MFOLD analyses of the possible DNA secondary structures (Zuker, 2003) indicated that the repeat sequence folds stably both as a single array and as two arrays (Fig. 3).

These observations suggest that the copy number of the repeat is highly variable, and turnover of the repeats is rapid. This, in turn, indicates that the same

copy number does not guarantee the same evolutionary history. Thus, we believe that copy number variation is not suitable as a genetic marker in *Lethenteron* species.

Regarding the high turnover rate, the case of jm2 is also worth noting. The repeat in jm2 has distinct sequences in its second to fourth arrays, which involve insertion of CTTTTT in the repeat (C1 and C2 in Fig. 2D). The sequence of the third repeat lost 5' nucleotide stretch of AATTGT (C3 in Fig. 2D). This replacement of the repeat sequence occurred without changing the number of repeats; jm1 retained the typical architecture of the arrays of *L. japonicum* (four B type arrays and three A type arrays). If slipped-strand mispairing accounts for the replacment of the repeat arrays, it should accompany the increase in copy number. Although we cannot exclude the possibility that the loss of the original B2 type of the repeat arrays occurred after the slippage event, we may need to consider other mechanisms of the gene conversion, such as inter-molecular recombination (Hoarau et al., 2002; Mjelle et al., 2008). Recent studies documented instances of mtDNA recombination (Hoarau et al., 2002; Mjelle et al., 2008), and some experimental evidence has been reported in other taxonomic groups (Lunt and Hyman, 1997; Ladoukakis and Zouros, 2001; Rokas et al., 2003).

In this regard, the case of the Kamo population of *L.* sp. S. also requires special attention. Two individuals, sk1 and sk2, possessed distinct repeat sequences

(Fig. 2E). The D1 sequence might have emerged via substitution and duplication of the internal sequence (Fig. 2E) and D3 might have emerged via further insertion and deletion of the sequence (Fig. 2F). Note that all of the repeat arrays were replaced by the D type, and no A type repeat arrays were left.

Characterization of the novel repeat sequences that emerged from tRNA genes

We found another type of the repeat sequence in some individuals of *L.* sp. S. from Senju, between $tRNA^{Thr}$ and $tRNA^{Glu}$ (tRNA-Thr/Glu repeat: Fig. 4). The novel repeat unit: tRNA-Thr/Glu repeat originated at the 3'-end of the $tRNA^{Thr}$ and the complementary sequence of the 3'-end of the $tRNA^{Glu}$ (Fig. 4B). The unit was repeated up to five times in *L.* sp. S. This expansion of the repeat is best explained by slipped-strand mispairing, facilitated by the secondary structure of the tRNA stems (Fig. 5). Additionally, the 3'-abutting sequence of the $tRNA^{Thr}$ anticodon and complementary sequence of $tRNA^{Glu}$ possess 13 bp of identical sequence (double ended arrows in Fig. 4B), which also facilitates the folding of the replication strand and subsequent slippage. The nucleotide sequence of the 13 bp nucleotide sequence was identical in *P. marinus* (Fig. 8; Lee and Kocher, 1995), although the repeat was not detected in the individual for which the full-length mtDNA was analyzed.

Complete nucleotide sequence of the mtDNA of L. *sp. S*

Because the repetitive sequence originated from the 3'-end of $tRNA^{Thr}$ and $tRNA^{Glu}$, we examined whether these tRNA genes were functional. We examined the entire sequence of mtDNA for *L.* sp. S (GenBank/EMBL/DDBJ Acc. No. AB565771). The gene arrangement was identical to that in *P. marinus* and *L. fluviatilis* (Lee and Kocher, 1995; Delarbre et al., 2000). There are several nucleotide sequences that need to be translated using $tRNA^{Thr}$ and $tRNA^{Glu}$, and no additional tRNA genes were found in the mitochondrial genome. Thus, we concluded that the two tRNA genes are likely functional.

Molecular markers for L. japonicum *and* L. kessleri

We compared the nucleotide sequences in the rest of the control region between *L. japonicum* and *L. kessleri*. Although there were some single nucleotide polymorphisms in the sequences, none was fixed in either species (data not shown). Searching for genetic markers, we also compared the approximately 400-bp sequences of the *SoxD* intron (Ohtani et al., 2008). Although there were two deletions, including one from a CA microsatellite, no fixed indel or substitution was observed in either

species (Sup. Fig. 1).

Our primary motivation for analyzing the control region of the *Lethenteron* species was to search for molecular markers suitable for species identification. However, our results suggest that the copy number is too variable for species identification, and we did not find reliable characteristics that distinguished *L. japonicum* from *L. kessleri.* Yamazaki and Goto (1998) found that only the MDH3 locus showed fixed allele differences in *L. japonicum* and *L. kessleri*. No allele was fixed in the remaining 26 allozymes tested and the *COI* sequence did not distinguish the two species (Yamazaki et al. 2006). No fixed differences were observed in the SoxD intron in the two species. This suggests that these two species diverged quite recently (Yamazaki and Goto, 1998, Yamazaki et al., 2006). Adult *L. japonicum* and *L. kessleri* are generally easy to distinguish because they differ markedly in size. Although *L. japonicum* continues to grow in the sea after metamorphosis, and *L. kessleri* stays in fresh water and does not grow after metamorphosis, Yamazaki et al. (1998) reported that some individuals of *L. japonicum* do stay in fresh water after metamorphosis. Additionally, it is difficult to distinguish the ammocoetes larvae of these species. Thus, for species identification, we need to find a reliable molecular marker. These two species may show fixed genetic markers in a restricted chromosome region around

MDH3 locus (Yamazaki and Goto, 1998).

Conclusions

In the present study, we found quite a dynamic nature of the repeat sequences of the *Lethenteron* mtDNA control region. Most of the 41 individuals examined show unique architectures of repeat arrays in their mtDNA control region. Our detail analyses of the repeat arrays provided evidences that most of the variations can be due to slipped-strand mispairing. Some cases, such as the repeat arrays found in specimens sk1, sk2 and jm2, might be better explained by gene conversion, due to intermolecular recombination. Although our primary purpose to find the molecular markers to distinguish between *L. japonicum* and *L. kessleri* was not achieved, the highly dynamic nature of the repeat sequences in *Lethenteron* mtDNA control region provided unique system to analyze the details of the molecular evolution of repeat sequences. In addition, the present study further strengthened the idea that these two species diverged quite recently. This rather encouraged us to seek for the evolutionary transition from the parasitic lifestyle of *L. japonicus* to non-parasitic lifestyle of *L. kessleri*. Investigations on the genomic sequences around the MDH3 locus may give us clues to understand the evolutionary transition.

Acknowledgement

We thank Nobuyuki Fuchigami for collecting specimens of *L.* sp. S. from Senju River.

We also thank Masanao Honda for his help in the network analyses.

References

Avise, J.C., 1991. Ten Unorthodox Perspectives on Evolution Prompted by Comparative Population Genetic Findings on Mitochondrial DNA. Annu. Rev. Genet. 25, 45-69.

Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A. and Saunders, N.C., 1987. Intraspecific Phylogeography: The Mitochondrial DNA Bridge Between Population Genetics and Systematics. Annu. Rev. Ecol. Syst. 18, 489-522.

Broughton, R.E. and Dowling, T.E., 1994. Length variation in mitochondrial DNA of the minnow *Cyprinella spiloptera*. Genetics 138, 179-190.

Buroker, N.E., Brown, J.R., Gilbert, T.A., O'Hara, P.J., Beckenbach, A.T., Thomas, W.K. and Smith, M.J., 1990. Length Heteroplasmy of Sturgeon Mitochondrial DNA: An Illegitimate Elongation Model. Genetics 124, 157-163.

Clement, M., Posada, D. and Crandall, K., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9, 1657-1660.

Delarbre, C., Escriva, H., Gallut, C., Barriel, V., Kourilsky, P., Janvier, P., Laudet, V. and Gachelin, G., 2000. The Complete Nucleotide Sequence of the Mitochondrial DNA of the Agnathan *Lampetra fluviatilis*: Bearings on the Phylogeny of Cyclostomes. Mol. Biol. Evol. 17, 519-529.

Hardisty, M.W. and Potter, I.C.: 1971, Paired species. In: Hardisty, M.W. and Potter, I.C. (Eds.), The Biology of Lampreys Vol. I. Academic Press, London.

Hoarau, G., Holla, S., Lescasse, R., Stan, W.T. and Olsen, J.L., 2002. Heteroplasmy and Evidence for Recombination in the Mitochondrial Control Region of the Flatfish Platichthys flesus. Mol. Biol. Evol. 19, 2261-2264.

Hubbs, C.I. and Potter, I.C.: 1971, Distribution, phylogeny and taxonomy. In: Hardisty, M.W. and Potter, I.C. (Eds.), The Biology of Lampreys Vol. I. Academic Press, London.

Ladoukakis, E.D. and Zouros, E., 2001. Direct evidence for homologous recombination in mussel (*Mytilus galloprovincialis*) mitochondrial DNA. Mol. Biol. Evol. 18, 1168-1175.

Lee, W. and Kocher, T.D., 1995. Complete sequence of a sea lamprey (Petromyzon marinus) mitochondrial genome: Early establishemnt of the vertebrate genome organization. Genetics 139, 873-887.

Ludwig, A., May, B., Debus, L. and Jenneckens, I., 2000. Heteroplasmy in the mtDNA control region of sturgeon. Genetics 156, 1933-1947.

Lunt, D.H. and Hyman, B.C., 1997. Animal mitochondrial DNA recombination. Nature 387, 247.

Mjelle, K.A., Karlsen, B.O., Jørgensen, T.E., Moum, T. and Johansen, S.D., 2008. alibut mitochondrial genomes contain extensive heteroplasmic tandem repeat arrays involved in DNA recombination. BMC Genomics 9, 10.

Mundy, N.I. and Helbig, A.J., 2004. Origin and Evolution of Tandem Repeats in the Mitochondrial DNA Control Region of Shrikes (*Lanius* spp.). J. Mol. Evol. 59, 250-257.

Ohtani, K., Yao, T., Kobayashi, M., Kusakabe, R., Kuratani, S. and Wada, H., 2008. Expression of Sox and fibrillar collagen genes in lamprey larval chondrogenesis with implications for the evolution of vertebrate cartilage. J. Exp. Zool. 310B, 596-607.

Potter, I.C., 1980. The Petromyzoniformes with Particular Reference to Paired Species. Can. J. Aquat. Sci. 37, 1595-1615.

Rokas, A., Ladoukakis, E. and Zouros, E., 2003. Animal mitochondrial DNA recombination revisted. Trends Ecol. Evol. 18, 411-417.

Vladykov, V.D. and Kott, E., 1979. Satellite species among the holarctic lampreys (Petromyzonidae). Can. J. Zool. 57, 860-867.

White, W.M. and Martin, H.R., 2009. Structure and Conservation of Tandem Repeats in the Mitochondrial DNA Control Region of the Least Brook Lamprey (*Lampetra aepyptera*). J. Mol. Evol. 68, 715-723.

Yamazaki, Y. and Goto, A., 1998. Genetic structure and differentiation of four Lethenteron taxa from the Far East, deduced from allozyme analysis. Env. Biol. Fish. 52, 149-161.

Yamazaki, Y., Sugiyama, H. and Goto, A., 1998. Mature dwarf males and females of the arctic lamprey, *Lethenteron japonicum*. Ichthyol. Res. 45, 404-408.

Yamazaki, Y., Yokoyama, R., Nishida, M. and Goto, A., 2006. Taxnomy and molecular phylogeny of *Lethenteron* lampreys in eastren Eurasia. J. Fish Biol. 68, 251-269.

Zanandrea, G., 1959. Specieation among lampreys. Nature 184, 380.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. . Nuc. Acids Res. 31, 3406-3415.

Figure legends

Figure 1. Nucleotide sequence of the non-coding region 2 (NC2) of the mtDNA control region of *Lethenteron Japonicum*, specimen js1. In this specimen six NC2 repeat is observed.

Figure 2. Sequence types and repeat architecture of the NC2 repeats. (A) Alignment of repeat sequences. (B) Network analysis of the repeat sequences. Numbers on the nodes indicate the nucleotide sites of the repeat in which substitutions occur between the sequences. (C) Schematic illustration of architecture of the repeat sequences in NC2 for each specimens. (D) Schematic illustration of a putative evolutionary history of the C-type array of repeats. Alignment of the C type repeat arrays. Evolutionary scheme for the C type. The C type of repeat array probably originated from the B2 type via 3' extension of the CTTTTT sequence. (E, F) Schematic illustration of a putative evolutionary history of the D-type array of repeats. Type D1 arose from type A3 via duplication of the 3' part of the repeat (E). Subsequently, D2 and D3 were derived from the D1 sequence (F).

Figure 3. Possible secondary structure of the NC2 repeat. (A) The secondary structure of the NC2 repeat of A3-5. The estimated free energy was -2.27 kcal at 120 mM Na^+, 2 mM Mg^{2+}, and 20°C. (B) The secondary structure of the two tandem repeats of A3-5. The estimated free energy was -9.79 kcal at 120 mM Na^+, 2 mM Mg^{2+}, and 20°C.

Figure 4. Sequence of the novel repeat originating from *tRNA^{Thr}* and *tRNA^{Glu}*. (A) Nucleotide sequence of the 3' half of the mtDNA control region of *L.* sp. S. (Ss3). This specimen possess five arrays of the tRNA-Thr/Glu repeat. (B) Sequence comparison of *tRNA^{Thr}*, *tRNA^{Glu}*, and the repeat sequence. 13 bp of identical nucleotide sequences are shown by the double-ended arrows. These sequences in *tRNA^{Thr}* and *tRNA^{Glu}* may have facilitated the slipped strand misparing (see text in detail). The tRNA anticodons are boxed. (C) Variation in the number of the repeats in five specimens of *L.* sp. S. White, Gray and Black box indicate the sequence of *tRNA^{Thr}* and $tRNA^{Glu}$ (complementary) and tRNA-Thr/Glu repeat as shown in (B).

Figure 5. Possible secondary structure of the sequence of the 3' half of *tRNA^{Thr}* and the 3'-half of the complementary strand of *tRNA^{Glu}*, which mediate the generation of the repeats. The estimated free energy was -6.13 kcal at 120 mM Na^+, 2 mM Mg^{2+}, and 20°C.

Figure 6. Alignment of the *tRNA^{Thr}* and *tRNA^{Glu}* in *Lethenteron* and *P. marinus*. The 13 stretches of the identical nucleotide sequence are conserved in *P. marinus*, although the repeat sequence was not reported.

Table1

Table 1. Collection sits and number of specimens examined.

Taxon	Collected Site	Number of specimens examined	Individual Recognition	Acc. Nos. for NC2	Acc Nos.for SoxD intron
Lethenteron japonicum	Shiribeshitoshibetsu R. (Hokkaido, Japan)	5	js1 ~ js5	AB567688-92	AB565490-92 (js1-3)
	Ishikari R. (Hokkaido, Japan)	3	ji1, ji2, ji3	AB567693-95	
	Mogami R. (Yamagata, Japan)	2	jm1, jm2	AB567696-97	
L. kessleri	Irtysh R. (Upper Ob, Kazakhstan)	5	k1 ~ k5	AB567698-702	AB565493-94 (k1-2)
L. sp. N	Shiribeshitoshibetsu R. (Hokkaido, Japan)	5	Ns1 ~ Ns5	AB567703-07	
	Kamo R. (Upper Shogawa, Toyama, Japan)	12	Nk1 ~ Nk12	AB567708-16, AB567729-31	
L. sp. S	Senju R. (Upper Onga, Fukuoka, Japan)	5	Ss1 ~ Ss5	AB567734-38	
	Kamo R. (Upper Shogawa, Toyama, Japan)	2	Sk1, Sk2	AB567732-33	

Table2

Table 2. Primers used in this study

	Primer name	Sequence
For NC2	NC2-F	5'-GCTGCCGAATACACAAAAACAACCATCAT-3'
	NC2-R	5'-TTGGCATGGAGGTTTCGTATAAGCCATCC-3'
For SoxD intron	SoxD-F	5'-GACGAGAGGCGGAAAATCCTTCAAGCTT-3'
	SoxD-R	5'-GAGGCTTGTACTTGTAGTCGGGATACTTC-3'
For complete mtDNA of *L.* sp. S.	Lsmt1F	5'-CTAGTAGATCTTCCTTCTCCTGCTAATAT-3'
	Lsmt1R	5'-GTTCCAGTGTAGGGTTAACGGTTATTAGTT-3'
	Lsmt2F	5'-CCCTATAACCACTATTAAGTAATCCTATAT-3'
	Lsmt2R	5'-GATCTTGTGCAATTTGAATAAGGAGAGTAA-3'
	Lsmt3F	5'-GCTCAAAGTGCAGGCTCTGCCACACTTCTT-3'
	Lsmt3R	5'-GGAGGGTAGCTAATCAGCTAAAAACTTTAA-3'
	Lsmt4F	5'-GGGATTATCTCTCACGTAGTTGCTTATTAT-3'
	Lsmt4R	5'-GTGATGGCCTAGAAAGGTGCCTTCTCGAA-3'
	Lsmt5F	5'-GTCTTTATTCTGCTTTTAACCCTCTACCTT-3'
	Lsmt5R	5'-CATGTTTAATGAAAAGAGAGCGGTTAAAGT-3'
	Lsmt6F	5'-CCTCTAATATCATTCTGATGACTTATAATA-3'
	Lsmt6R	5'-CCAGTGCTTTATATTTAAGCTATCAAAGCT-3'

Figure1

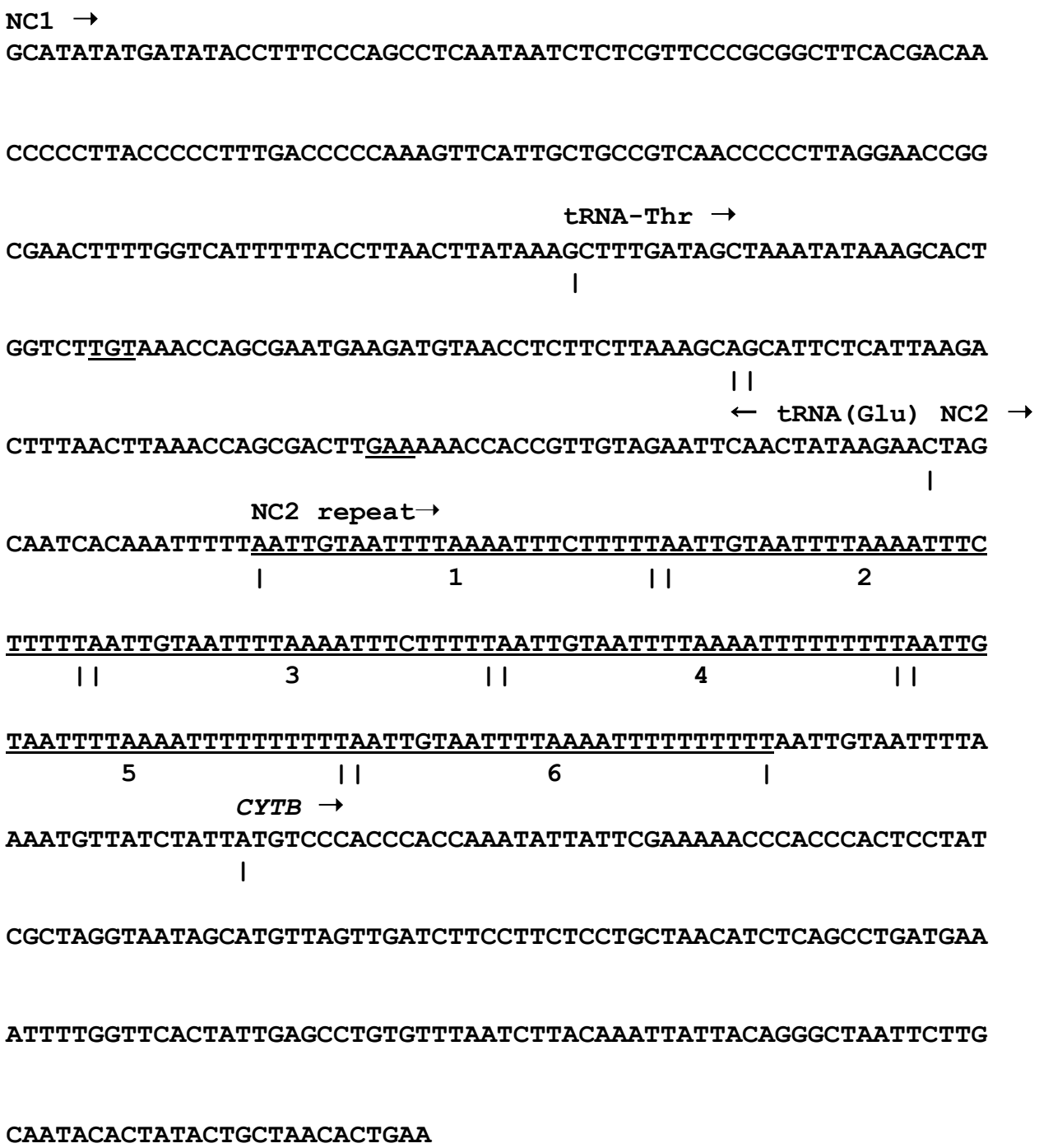

NC1 →
GCATATATGATATACCTTTCCCAGCCTCAATAATCTCTCGTTCCCGCGGCTTCACGACAA
CCCCCTTACCCCCTTTGACCCCCAAAGTTCATTGCTGCCGTCAACCCCCTTAGGAACCGG
tRNA-Thr →
CGAACTTTTGGTCATTTTTACCTTAACTTATAAAGCTTTGATAGCTAAATATAAAGCACT
GGTCTTGTAAACCAGCGAATGAAGATGTAACCTCTTCTTAAAGCAGCATTCTCATTAAGA
← tRNA(Glu) NC2 →
CTTTAACTTAAACCAGCGACTTGAAAAACCACCGTTGTAGAATTCAACTATAAGAACTAG
NC2 repeat→
CAATCACAAATTTTTAATTGTAATTTTAAAATTTCTTTTTAATTGTAATTTTAAAATTTC
1
2
TTTTTAATTGTAATTTTAAAATTTCTTTTTAATTGTAATTTTAAAATTTTTTTTTAATTG
3
4
TAATTTTAAAATTTTTTTTTTAATTGTAATTTTAAAATTTTTTTTTAATTGTAATTTTA
5
6
CYTB →
AAATGTTATCTATTATGTCCCACCCACCAAATATTATTCGAAAAACCCACCCACTCCTAT
CGCTAGGTAATAGCATGTTAGTTGATCTTCCTTCTCCTGCTAACATCTCAGCCTGATGAA
ATTTTGGTTCACTATTGAGCCTGTGTTTAATCTTACAAATTATTACAGGGCTAATTCTTG
CAATACACTATACTGCTAACACTGAA

Figure2A

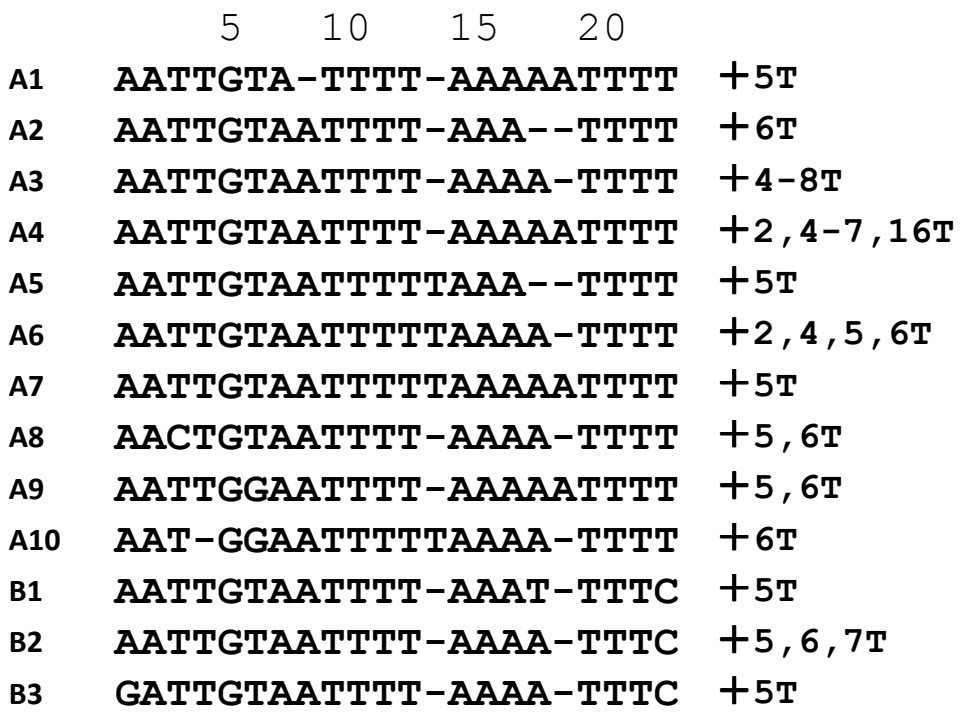

5 10 15 20
A1 AATTGTA-TTTT-AAAAATTTT +5T
A2 AATTGTAATTTT-AAA--TTTT +6T
A3 AATTGTAATTTT-AAAA-TTTT +4-8T
A4 AATTGTAATTTT-AAAAATTTT +2,4-7,16T
A5 AATTGTAATTTTTAAA--TTTT +5T
A6 AATTGTAATTTTTAAAA-TTTT +2,4,5,6T
A7 AATTGTAATTTTTAAAAATTTT +5T
A8 AACTGTAATTTT-AAAA-TTTT +5,6T
A9 AATTGGAATTTT-AAAAATTTT +5,6T
A10 AAT-GGAATTTTTAAAA-TTTT +6T
B1 AATTGTAATTTT-AAAT-TTTC +5T
B2 AATTGTAATTTT-AAAA-TTTC +5,6,7T
B3 GATTGTAATTTT-AAAA-TTTC +5T

Figure2B

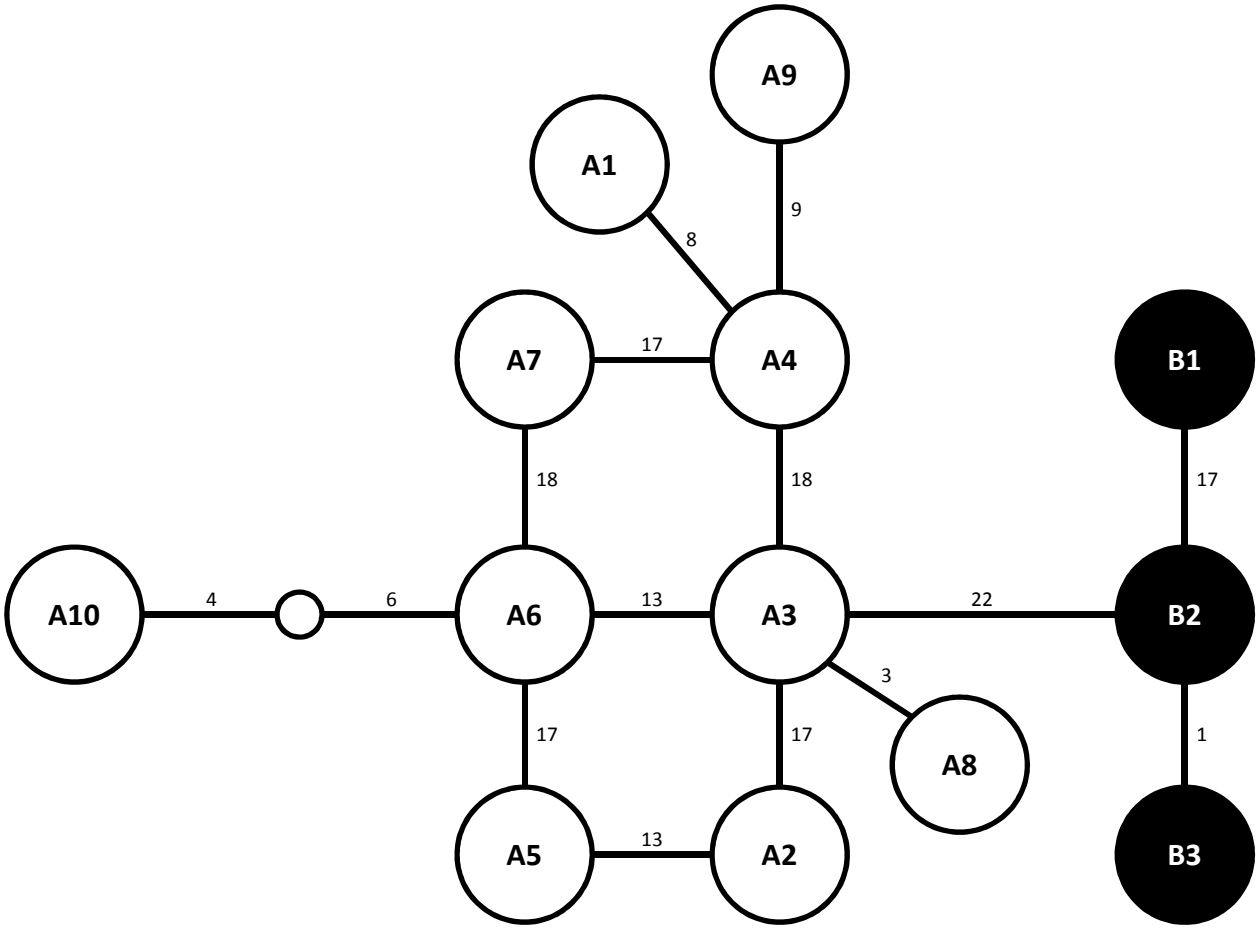

A9
A1
9
8
A7
17
A4
B1
18
18
17
A10
4
6
A6
13
A3
22
B2
3
17
17
1
A8
A5
13
A2
B3

Figure2C

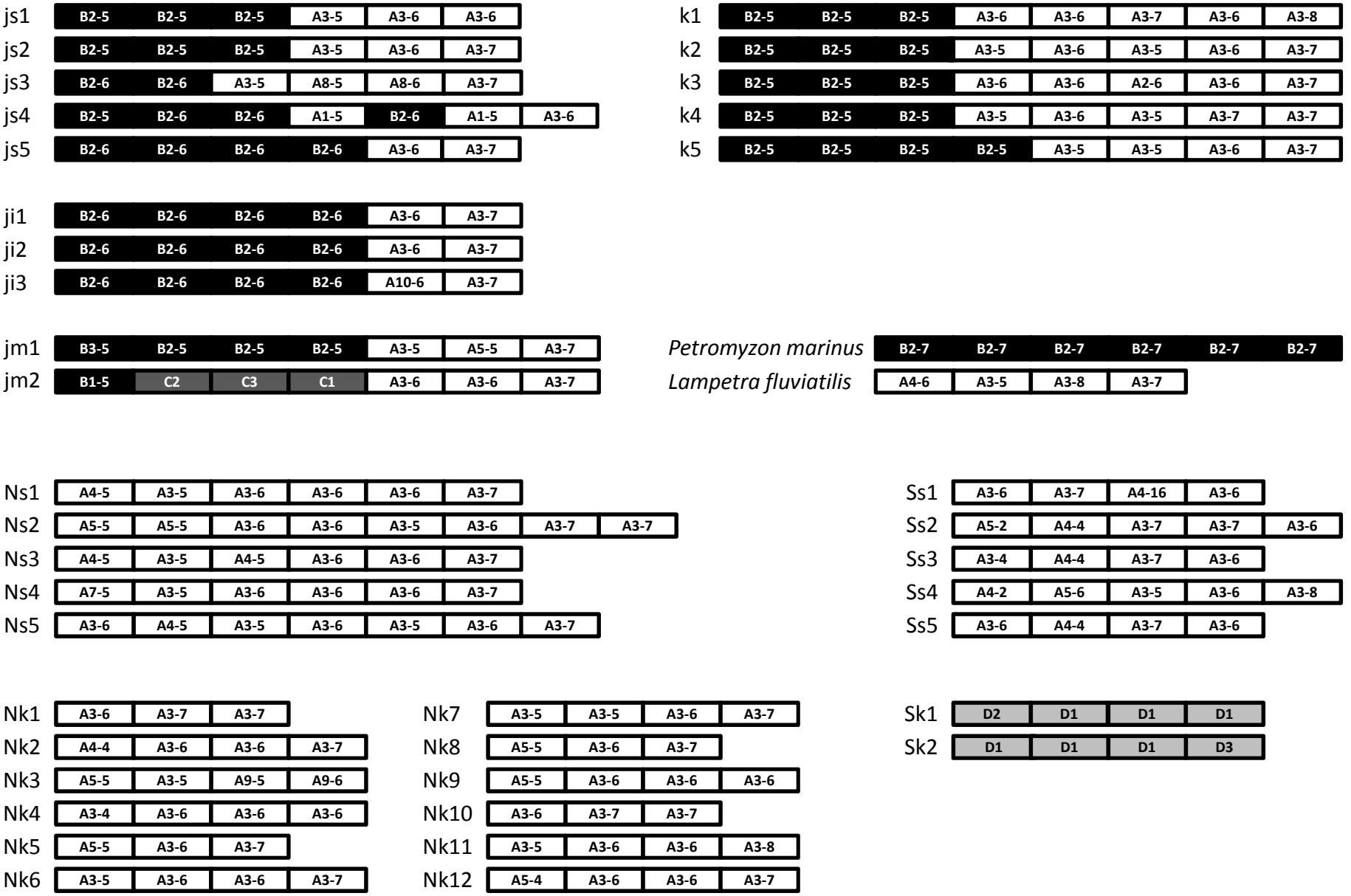

js1 B2-5 B2-5 B2-5 A3-5 A3-6 A3-6
js2 B2-5 B2-5 B2-5 A3-5 A3-6 A3-7
js3 B2-6 B2-6 A3-5 A8-5 A8-6 A3-7
js4 B2-5 B2-6 B2-6 A1-5 B2-6 A1-5 A3-6
js5 B2-6 B2-6 B2-6 B2-6 A3-6 A3-7
k1 B2-5 B2-5 B2-5 A3-6 A3-6 A3-7 A3-6 A3-8
k2 B2-5 B2-5 B2-5 A3-5 A3-6 A3-5 A3-6 A3-7
k3 B2-5 B2-5 B2-5 A3-6 A3-6 A2-6 A3-6 A3-7
k4 B2-5 B2-5 B2-5 A3-5 A3-6 A3-5 A3-7 A3-7
k5 B2-5 B2-5 B2-5 B2-5 A3-5 A3-5 A3-6 A3-7
ji1 B2-6 B2-6 B2-6 B2-6 A3-6 A3-7
ji2 B2-6 B2-6 B2-6 B2-6 A3-6 A3-7
ji3 B2-6 B2-6 B2-6 B2-6 A10-6 A3-7
jm1 B3-5 B2-5 B2-5 B2-5 A3-5 A5-5 A3-7
jm2 B1-5 C2 C3 C1 A3-6 A3-6 A3-7
Petromyzon marinus B2-7 B2-7 B2-7 B2-7 B2-7 B2-7
Lampetra fluviatilis A4-6 A3-5 A3-8 A3-7
Ns1 A4-5 A3-5 A3-6 A3-6 A3-6 A3-7
Ns2 A5-5 A5-5 A3-6 A3-6 A3-5 A3-6 A3-7 A3-7
Ns3 A4-5 A3-5 A4-5 A3-6 A3-6 A3-7
Ns4 A7-5 A3-5 A3-6 A3-6 A3-6 A3-7
Ns5 A3-6 A4-5 A3-5 A3-6 A3-5 A3-6 A3-7
Ss1 A3-6 A3-7 A4-16 A3-6
Ss2 A5-2 A4-4 A3-7 A3-7 A3-6
Ss3 A3-4 A4-4 A3-7 A3-6
Ss4 A4-2 A5-6 A3-5 A3-6 A3-8
Ss5 A3-6 A4-4 A3-7 A3-6
Nk1 A3-6 A3-7 A3-7
Nk2 A4-4 A3-6 A3-6 A3-7
Nk3 A5-5 A3-5 A9-5 A9-6
Nk4 A3-4 A3-6 A3-6 A3-6
Nk5 A5-5 A3-6 A3-7
Nk6 A3-5 A3-6 A3-6 A3-7
Nk7 A3-5 A3-5 A3-6 A3-7
Nk8 A5-5 A3-6 A3-7
Nk9 A5-5 A3-6 A3-6 A3-6
Nk10 A3-6 A3-7 A3-7
Nk11 A3-5 A3-6 A3-6 A3-8
Nk12 A5-4 A3-6 A3-6 A3-7
Sk1 D2 D1 D1 D1
Sk2 D1 D1 D1 D3

Figure2D

```
          5    10    15    20    25    30
B2    AATTGTAATTTTAAAATTTC----------
C1    AATTGTAATTTTAAAATTTCTTTTTCTTTTT
C2    AATTGTAATTTTAAATTTTCTTTTTCTTTTT
C3    ------AATTTTAAATTTTCTTTTTCTTTTT
```

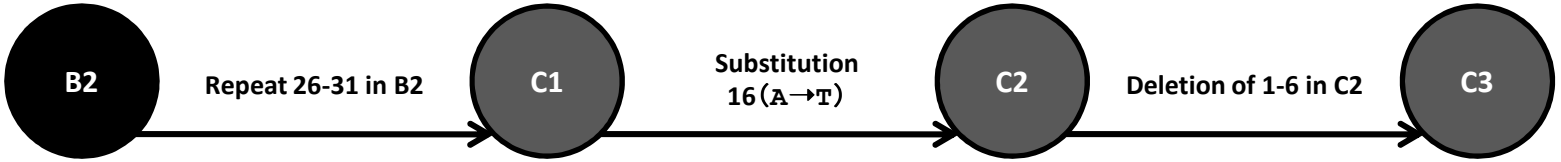


Figure2E

```
                5    10    15    20    25    30    35    40    45    50
A3 x2      AATTGTAATTTTAAAATTTTTTTTTTAATTGTAATTTTAAAATTTTTTTTTT
D1         AATTTTAATTTTAAAATTTTATTTTT-------ATTTTAAAAATTTTTTTTT
```

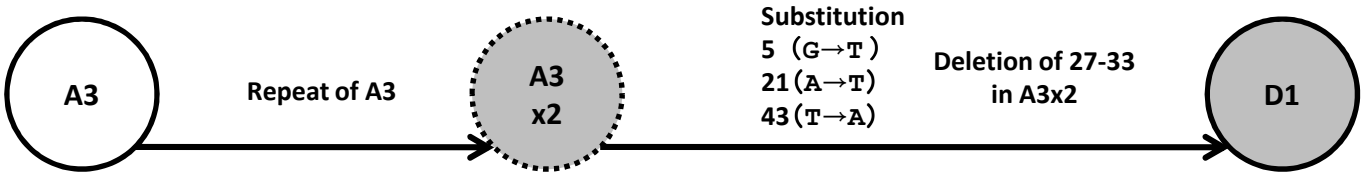


Figure2F

```
          5    10    15    20    25    30    35    40    45    50    55
D1   AATTTTAATTTTAAAATTTTATTTTTATTTT-----------AAAAATTTTTTTTT-
D2   AATTTTAATTTTAAAATTTT-TTTTTATTTT-----------AAAAATTTTTTTTT-
D3   ------AATTTTAAAATTTTATTTTTATTTTTATTTTTATTTTAAAAATTTTTTTTT
```

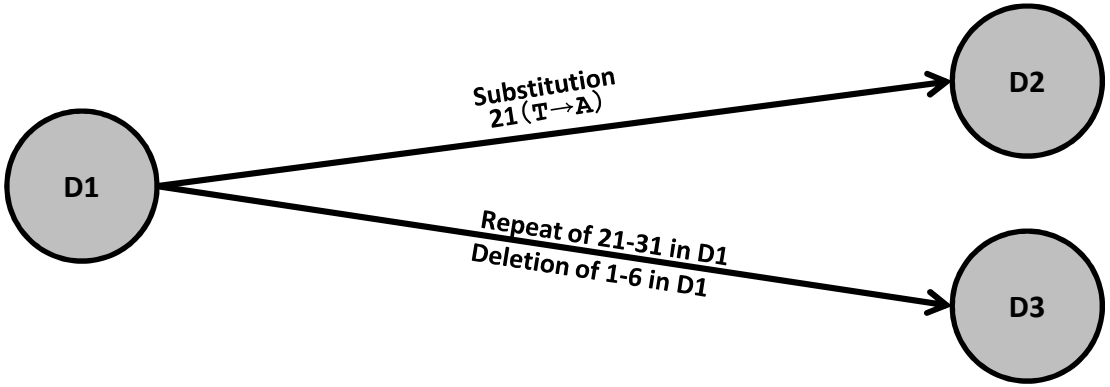


Figure3A

Figure3B

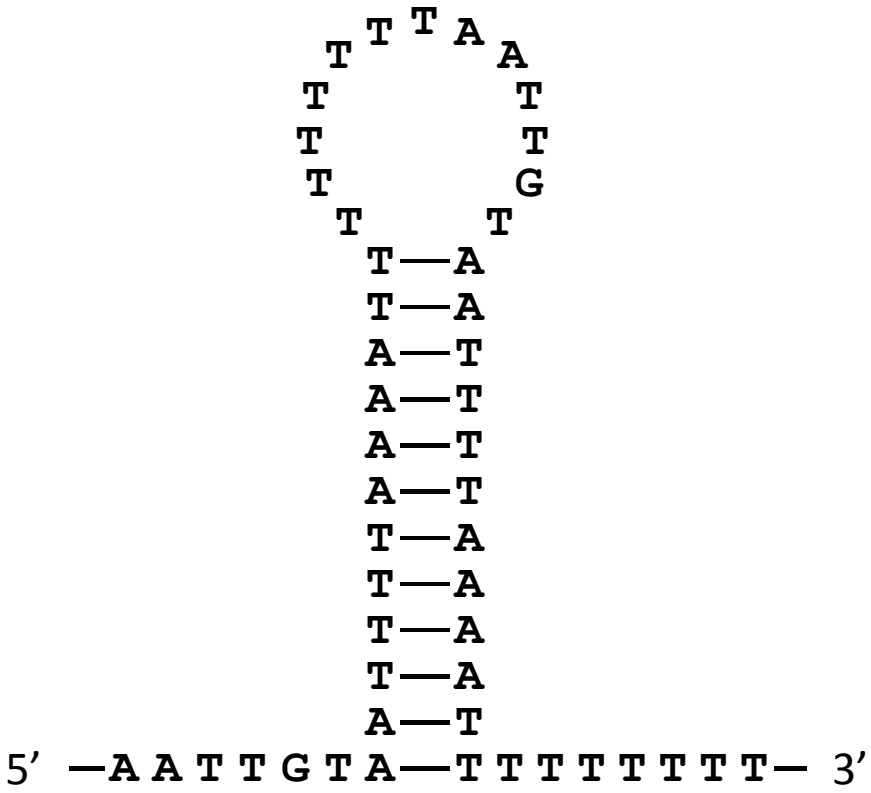

T T A
T A
T T
T T
T G
T T
T—A
T—A
A—T
A—T
A—T
A—T
T—A
T—A
T—A
T—A
A—T
5’ —A A T T G T A—T T T T T T T T T— 3’

Figure4A

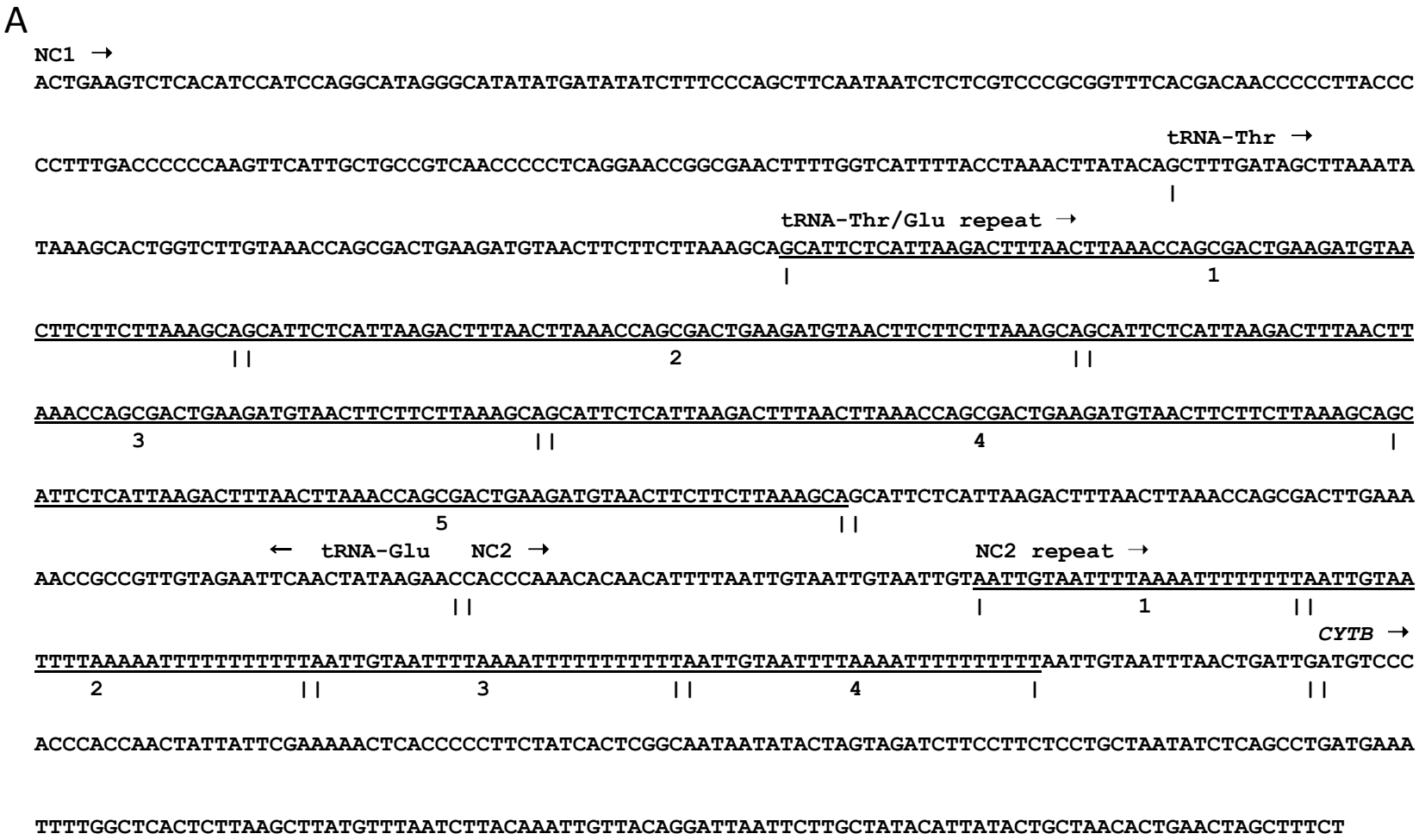

A
NC1 →
ACTGAAGTCTCACATCCATCCAGGCATAGGGCATATATGATATATCTTTCCCAGCTTCAATAATCTCTCGTCCCGCGGTTTCACGACAACCCCCTTACCC
tRNA-Thr →
CCTTTGACCCCCCAAGTTCATTGCTGCCGTCAACCCCCTCAGGAACCGGCGAACTTTTGGTCATTTTACCTAAACTTATACAGCTTTGATAGCTTAAATA
tRNA-Thr/Glu repeat →
TAAAGCACTGGTCTTGTAAACCAGCGACTGAAGATGTAACTTCTTCTTAAAGCAGCATTCTCATTAAGACTTTAACTTAAACCAGCGACTGAAGATGTAA
1
CTTCTTCTTAAAGCAGCATTCTCATTAAGACTTTAACTTAAACCAGCGACTGAAGATGTAACTTCTTCTTAAAGCAGCATTCTCATTAAGACTTTAACTT
2
AAACCAGCGACTGAAGATGTAACTTCTTCTTAAAGCAGCATTCTCATTAAGACTTTAACTTAAACCAGCGACTGAAGATGTAACTTCTTCTTAAAGCAGC
3
4
ATTCTCATTAAGACTTTAACTTAAACCAGCGACTGAAGATGTAACTTCTTCTTAAAGCAGCATTCTCATTAAGACTTTAACTTAAACCAGCGACTTGAAA
5
← tRNA-Glu NC2 →
NC2 repeat →
AACCGCCGTTGTAGAATTCAACTATAAGAACCACCCAAACACAACATTTTAATTGTAATTGTAATTGTAATTGTAAATTTTAAAATTTTTTTTAATTGTAA
1
CYTB →
TTTTAAAAATTTTTTTTTTTAATTGTAATTTTAAAATTTTTTTTTTTAATTGTAATTTTAAAATTTTTTTTTTAATTGTAATTTAACTGATTGATGTCCC
2
3
4
ACCCACCAACTATTATTCGAAAAACTCACCCCCTTCTATCACTCGGCAATAATATACTAGTAGATCTTCCTTCTCCTGCTAATATCTCAGCCTGATGAAA
TTTTGGCTCACTCTTAAGCTTATGTTTAATCTTACAAATTGTTACAGGATTAATTCTTGCTATACATTATACTGCTAACACTGAACTAGCTTTCT

Figure4B

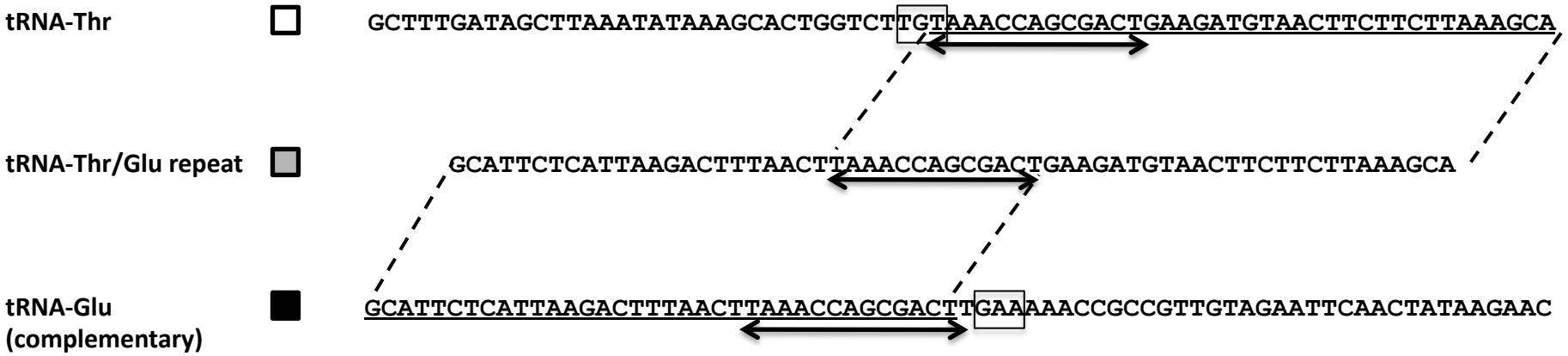

tRNA-Thr
GCTTTGATAGCTTAAATATAAAGCACTGGTCTTGTAAACCAGCGACTGAAGATGTAACTTCTTCTTAAAGCA
tRNA-Thr/Glu repeat
GCATTCTCATTAAGACTTTAACTTAAACCAGCGACTGAAGATGTAACTTCTTCTTAAAGCA
tRNA-Glu
(complementary)
GCATTCTCATTAAGACTTTAACTTAAACCAGCGACTTGAAAAACCGCCGTTGTAGAATTCAACTATAAGAAC

Figure4C

Figure5

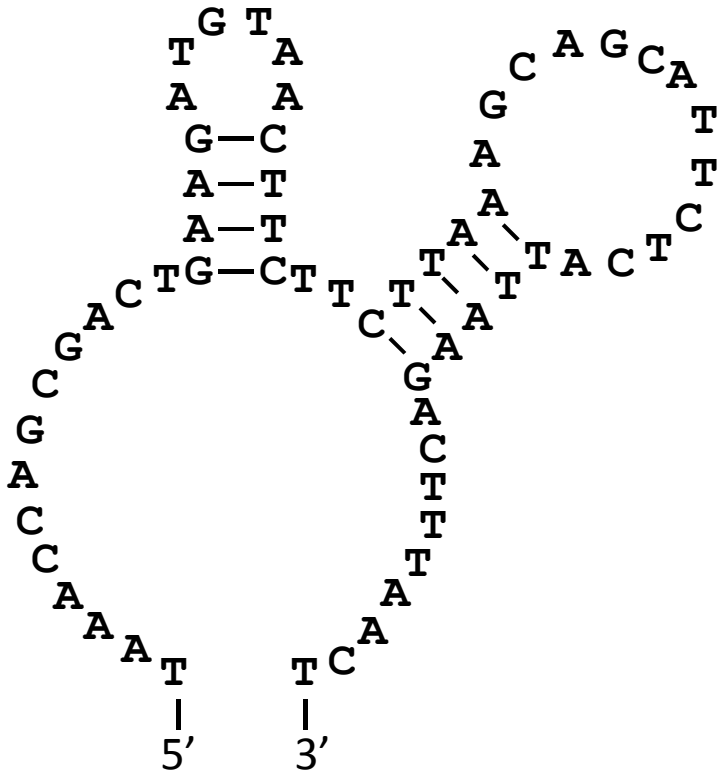
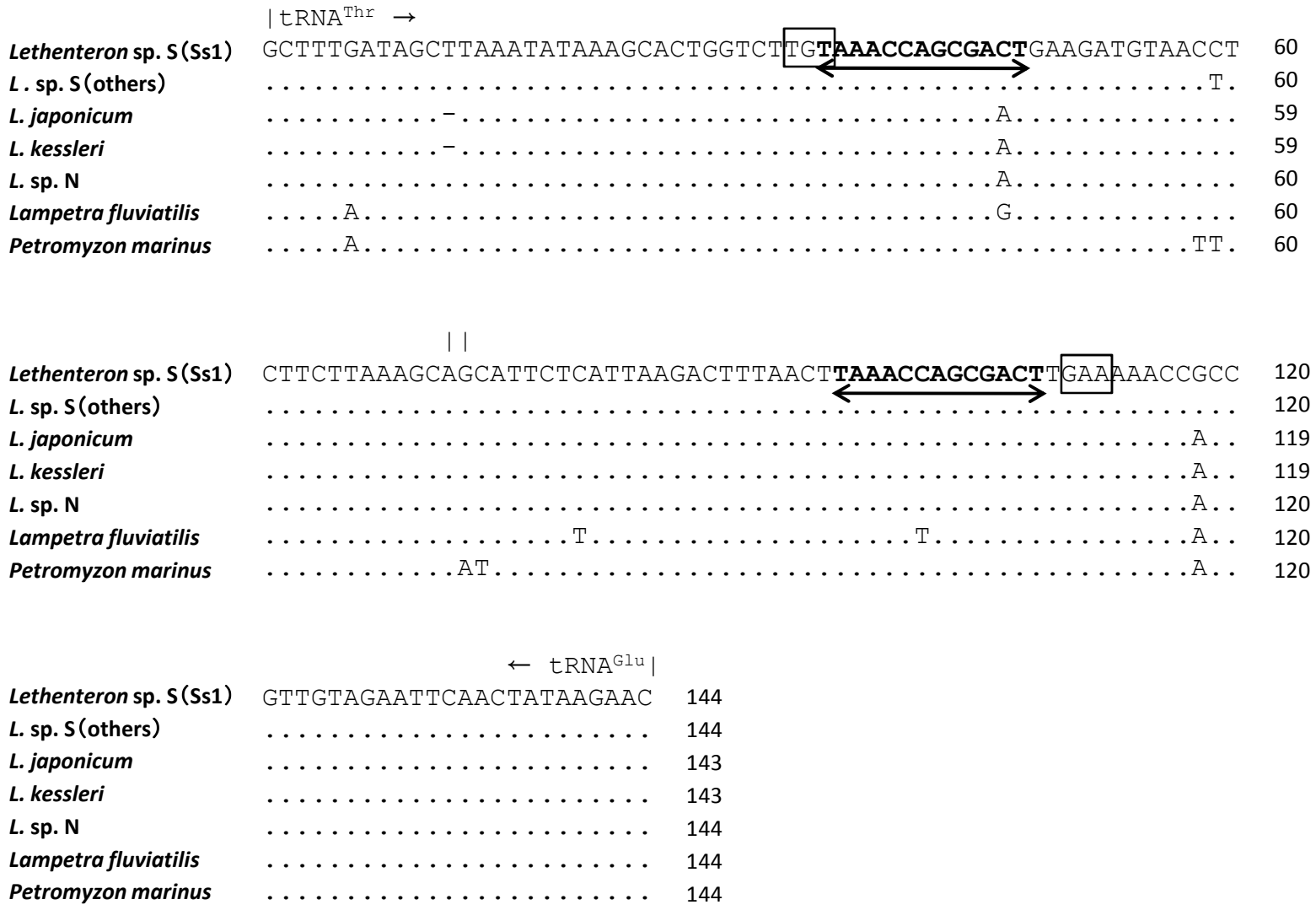


Figure6


tRNA^{Thr}, tRNA^{Glu}
|tRNA^{Thr} →
Lethenteron sp. S(Ss1) GCTTTGATAGCTTAAATATAAAGCACTGGTCTTGTTAAACCAGCGACTGAAGATGTAACCT 60
L. sp. S(others) .....T. 60
L. japonicum .....-.....A..... 59
L. kessleri .....-.....A..... 59
L. sp. N .....A..... 60
Lampetra fluviatilis .....A.....G..... 60
Petromyzon marinus .....A.....TT. 60
||
Lethenteron sp. S(Ss1) CTTCTTAAAGCAGCATTCTCATTAAGACTTTAACTTAAACCAGCGACTTGAAAAACCGCC 120
L. sp. S(others) 120
L. japonicum .....A.. 119
L. kessleri .....A.. 119
L. sp. N .....A.. 120
Lampetra fluviatilis .....T.....T.....A.. 120
Petromyzon marinus .....AT.....A.. 120
← tRNA^{Glu}|
Lethenteron sp. S(Ss1) GTTGTAGAATTCAACTATAAGAAC 144
L. sp. S(others) 144
L. japonicum 143
L. kessleri 143
L. sp. N 144
Lampetra fluviatilis 144
Petromyzon marinus 144

SFig1

Click here to download Supplementary material for on-line publication only: SFig1.pdf