Journal of Fish Biology (1998) **53,** 179–191
Article No. jb980692

Sequence length and variation in the mitochondrial control region of two freshwater gobiid fishes belonging to *Rhinogobius* (Teleostei: Gobioidei)

I-S. CHEN*†, C.-H. HSU†, C.-F. HUI‡, K.-T. SHAO‡, P. J. MILLER*§ AND L.-S. FANG†

**School of Biological Sciences, University of Bristol, Bristol, BS8 1UG, U.K.; †Institute of Marine Resources, National Sun Yat-Sen University, Kaohsiung, Taiwan, R.O.C. and ‡Institute of Zoology, Academia Sinica, Taipei, Taiwan, R.O.C.*

(*Received 5 April 1997, Accepted 17 March 1998*)

The control region (D-loop) of mitochondrial DNA (mtDNA) was amplified and sequenced for eight samples of the rhinogobies *Rhinogobius maculafasciatus* and *R. giurinus* from Taiwan and southern China. The control regions of both species are of 841–842 bp; the length of these sequences being the most compact among all known sequences in teleost fishes. Three conserved sequence blocks (CSB) were observed. The full D-loop and $tRNA^{Phe}$ gene sequences were determined and compared with other fishes. The interspecific sequence divergence between the two species is 11·3–11·7%; and the intraspecific variation in *R. guirinus* 0·8–1·8%. Results suggest that the control region of *Rhinogobius* is informative for phylogenetic reconstruction at both intraspecific and interspecific levels in this gobiid genus.

Key words: *Rhinogobius*; Gobiidae; mitochondrial genome; mtDNA sequence; D-loop; tRNA; polymerase chain reaction.

INTRODUCTION

Mitochondrial DNA (mtDNA) has been used widely in the study of evolutionary problems and population structure because of its rapid evolutionary rate and almost complete maternal inheritance (Brown *et al.*, 1979; Wilson *et al.*, 1985). Recently, the application of nucleotide sequence data in phyletic reconstruction has risen in popularity due to the potentially high resolving power of this technique (Avise *et al.*, 1988; Sang *et al.*, 1994; Yang *et al.*, 1994).

Rhinogobius Gill, 1859 (type *R. similis* Gill, 1859) is a genus of common freshwater gobiid fishes occurring in eastern Asia, and contains both anadromous and landlocked species (Mizuno, 1960; Mizuno & Goto, 1989; Iguchi & Mizuno, 1990; Chen & Shao, 1996). *Rhinogobius* has been regarded as the most speciose genus of freshwater gobies in Western Pacific drainages, with at least 40 species in a wide generic distribution from the Amur Basin in Russia to the Me-Nan Basin in Thailand and also on the islands of Japan, Taiwan and the Philippines (Mizuno & Goto, 1989; Chen & Shao, 1996; Chen & Miller, unpubl. data). The phylogeny of *Rhinogobius* species is difficult to resolve by morphological analysis because meristic characters have wide overlap, although specific coloration patterns may be useful especially in the *R. brunneus* complex

§Author to whom correspondence should be addressed. Email: peter.miller@bristol.ac.uk

0022–1112/98/070179+13 $30.00/0

TABLE I. Sample localities of two *Rhinogobius* species

Species	Code	Locality	Date
R. maculafasciatus	RM-T1, T2	Kaoping River, Pingtung, Taiwan	Dec. 1992
R. giurinus	RG-T1	Darhu, Juan River, Iran, Taiwan	Jul. 1993
	RG-F1	Jiulung River, Fujian, China	Aug. 1994
	RG-Y1, Y2, Y3, Y4	Nanpan River, Kuming, Yunan, China	Jul. 1996

(Mizuoka, 1974; Hayashi in Masuda *et al.*, 1984; Akihito *et al.*, 1993). To investigate the phylogeny of the rapidly evolving *Rhinogobius* species by molecular sequence data, it was decided to amplify and sequence the mitochondrial control region (also called D-loop, Brown, 1983) using the polymerase chain reaction (PCR).

Although several sequences of mtDNA control regions of fishes have been employed for genomic research and phylogenetic studies on inter- or intraspecfic groups (Shedlock *et al.*, 1992; Tzeng *et al.*, 1992; Chang *et al.*, 1994; Lee *et al.*, 1995; Jean *et al.*, 1995; Cesaroni *et al.*, 1997), the control region of no gobiid fish has been published. The aim of the present paper is not only to provide a first report on the complete sequences of gobiid mitochondrial control regions from two *Rhinogobius* species collected from Taiwan and China, but also to compare interspecific differentiation and interspecific variation among these two species.

MATERIALS AND METHODS

SAMPLE COLLECTION AND DNA EXTRACTION

The specimens of two *Rhinogobius* species were collected by hand-net during snorkeling or electrofishing over the period 1992–1996. Two individuals of the Spot-banded Rhinogoby, *Rhinogobius maculafasciatus* Chen & Shao, 1996, were collected from Taiwan and six individuals of the Paradise Rhinogoby, *Rhinogobius giurinus* (Rutter, 1897), were collected from both Taiwan and southern China (Table I). Some specimens were kept at -20° C until the extraction of DNA, while others were fixed in 95% ethanol and transferred subsequently to the freezer. The method of crude DNA extraction followed Sambrook *et al.* (1989).

PCR AMPLIFICATION AND SEQUENCING OF MTDNA

A fragment of mtDNA that included the D-loop region, $tRNA^{Phe}$ gene and partial 12 S rRNA gene was amplified by a pair of primers, Pa (5′-CTTACTATCAACTAAAGC-3′) and Pb (5′-GGGCCATTCTC ACGGGGATGCG-3′). These and three other primers Pc (5′-CCTGAAAACCCCCCGGAAAC-3′), Pd (5′-AATTAAGCGATGCGACCTCG-3′), Pe (5′-TCGAGAGTTTCCTGTTTCCG-3′) were used for DNA sequencing. Pa and Pb were designed by comparison with sequences for other perciforms from Genbank and from Chen *et al.* (unpubl. data). Primers Pc, Pd, Pe were designed from conserved sequences of *Rhinogobius* obtained from the DNA sequencing with primers Pa and Pb.

Polymerase chain reactions (PCR) were performed in 100-µl reaction volumes. Each sample contained 10 µl of 10 × reaction buffer (10 mM Tris–HCl, pH 9·0; 50 mM $MgCl_2$; 0·1% (w/v) gelatin; 1% Triton X-100), 0·4 µM of each primer, 0·2 µM of each dNTP, 1 µl of DNA extract, and 2–2·5 units of super-*Taq* polymerase (HT Biotechnology). PCR amplifications were carried out in a Perkin Elmer-Cetus thermal cycler using the

following cycling parameters: one cycle of pre-denaturation at 94° C for 1 min; 30–40 cycles of denaturation at 94° C for 1 min, primer annealling at 46–50° C for 1 min, and extension at 72° C for 2–2·5 min.

The double-stranded DNA product was purified by electroelution (Sambrook *et al.*, 1989), or using the Genclean II Kit (Bio 101). Vacuum-dried DNA was resuspended in 20 µl of distilled H_2O.

DNA sequencing was performed with the dideoxy-nucleotide chain termination method (Sanger *et al.*, 1977) using a CircumVent Thermal Cycle Dideoxy DNA Sequencing Kit (New England Biolabs Inc.), labelling with [a^{35}S]-dATP. Reactions were carried out with the following cycling parameters: 20 cycles of denaturation at 95° C for 30 s, primer annealling at 50–55° C for 30 s, and extension at 72° C for 30 s. Completed sequencing reactions were electrophoresed in 6% polyacrylamide/7 M urea gel. Subsequently, gels were fixed, dried and autoradiographed on X-ray negative for 24–96 h. Alternatively, automated sequencing was performed as recommended with a PRISM™ Ready Reaction DyeDeoxy™ Terminator Cycle Sequencing Kit for an ABI Model 373 DNA sequencer.

SEQUENCE ANALYSIS

MtDNA sequences enclosing a part of $tRNA^{Pro}$ gene, the control region (D-loop), the $tRNA^{Phe}$ gene, and a part of the 12 S rRNA gene were aligned by using the PILEUP program of GCG software package (Genetic Computer Group, version 7.0; Devereux *et al.*, 1991). After minor adjustments, they were compared with published sequences for mammals (Anderson *et al.*, 1981; Bibb *et al.*, 1981) and fishes such as Acipenseriformes (Buroker *et al.*, 1990), Clupeiformes (Lee *et al.*, 1995), Salmoniformes (Digby *et al.*, 1992; Shedlock *et al.*, 1992), Cypriniformes (Tzeng *et al.*, 1992; Chang *et al.*, 1994), Gadiformes (Johansen *et al.*, 1990; Lee *et al.*, 1995), Perciformes (Cecconi *et al.*, 1995; Jean *et al.*, 1995; Lee *et al.*, 1995), and Pleuronectiformes (Lee *et al.*, 1995) to verify boundaries and the alignment of genes. Two techniques were used to construct phylogenetic relationships, distance (UPGMA) and maximum parsimony analysis. The UPGMA clustering was performed using PHYLIP 3.5 (Felsenstein, 1993) and distances were calculated by the Kimura two-parameter model/algorithm in DNADIST. Heuristic parsimony analysis with 1000 bootstrap replications was conducted using parsimony software PAUP 3.1.1 (Swofford, 1993).

RESULTS

DNA LENGTH AND SEQUENCE

For each specimen, the complete control region and $tRNA^{Pro}$ gene was amplified and both strands were sequenced (Fig. 1). The length of D-loop sequences for *Rhinogobius maculafasciatus* is 841 bp and varies from 841 bp (five individuals) to 842 bp (one individual) in *R. giurinus*. The gobiid *Rhinogobius* D-loop sequence seems to be the most compact of all the fully known sequences for teleost fishes (Table II).

The mitochondrial $tRNA^{Phe}$ gene from the two *Rhinogobius* species is identical except that a single substitution has occurred in the right, *TΨC* stem. The length of $tRNA^{Phe}$ is equal to that of the rainbow trout *Oncorhynchus mykiss* (Walbaum) (68 bp) (Digby *et al.*, 1992), but slightly shorter than that of the carp *Cyprinus carpio* L. (69 bp) (Chang *et al.*, 1994) and a loach *Crossostoma lacustre* Steindachner (70 bp) (Tzeng *et al.*, 1992).

DIVERGENCE RATES IN DIFFERENT REGIONS

The overall sequence divergence between the two *Rhinogobius* species ranges from 11·3 to 11·7% (Table III). The percentages of divergence per 50 bp of

```
        tRNA^Pro  D-loop
          <---+---->
            1          TAS-1                                              60
RM-T1,2 CTTTGTAAAG ACATATATGT ATTAACACCA TATATTTATG TTAACCATAT CAATAATGTA
RG-T1   .........T .......... ...T...... .......... .......... A.........
RG-F1   .........T .......... ...T...... .......... .......... A.........
RG-Y1   .........T .......... ...T...... .......... .......... A.........
RG-Y2,3 .........T .......... ...T...... .......... .......... A.........
RG-Y4   .........T .......... ...T...... .......... .......... A.........

           61          TAS-2                                              120
RM-T1,2 TTAGGACATC ATCTATGTAT AGTACTCATT ACTCGCTTTT GTCCATTCAT ACATAACCAT
RG-T1   .........T G-........ .ACC...... .A.A.GC... .......... ....C.....
RG-F1   .........T .-........ .ACC...... .A.A.GC... .......... ....C.....
RG-Y1   .........T .-........ .A.C...... CA.A.GC... .......... ...CC.....
RG-Y2,3 .........T .-........ .A.C...... CA.A.GC... .......... ....C.....
RG-Y4   .........T .-........ .A.C...... CA.A.GC... .......... ....C.....

        121                                                                180
RM-T1,2 GCTTTCAAAT TTCAACAGAA TGCA--TGAA ATAAAGTTGA GCACAACTGC ATACTAATAG
RG-T1   T.......GA C.....TA.. .A..TT.A.C .C.G.AA.T. -...T..... .C........
RG-F1   T.......GA C.....TA.. .A..TT.A.C .C.G.AA.T. -...T..... .C........
RG-Y1   T.....T.GA C.....TAT. .A..TT.A.C .C.G.AA.T. -...T..... .C........
RG-Y2,3 T.......GA C.....TA.. .A..TT.A.C .C.G.AA.T. -...T..... .C........
RG-Y4   T.......GA A.....TA.. .A..TT.A.C .C.G.AA.T. -...T..... .C........

        181                                                                240
RM-T1,2 GAATTACCCA ACTGTGTGAC TCACAACAAG TTAAGCCCTA ACACAGACTT TAGGTACTCA
RG-T1   A.......... .......A.. C..T.T...A C.......... .......AG. ..A.......
RG-F1   A.......... .....A.A.. ...T.T...A C.......... .......AG. ..A.......
RG-Y1   A.......... .......A.. ...T.T...A C.......... .......AG. ..A.......
RG-Y2,3 A.......... .......A.. ...T.T...A C.......... .......AG. ..A.......
RG-Y4   A.......... .......A.. ...T.T...A C.......... .......AG. ..A.......

        241                                                                300
RM-T1,2 TTTTATCCAT TACTCACCTT CTCTGCAAGT CAATATCGGA TGCAGACAAG AATTGCATTT
RG-T1   .......... .......... ....T..... .......... ......T... ..........
RG-F1   .......... .......... ....T..... .......... ......T... ...A......
RG-Y1   .......... .......... ....T..... .......... .......... ...A......
RG-Y2,3 .......... .......... ....T..... .......... .......... ...A......
RG-Y4   .......... .......... ....T..... .......... .......... ...A......

        301                                                                360
RM-T1,2 TAATTAAGCG ATGCGACCTC GGTTAACGAA GGTGAGGGAC AAGTATTCGT GGGGGTTTCA
RG-T1   ....G..... .....TTA.. .......... .......... .......... ..........
RG-F1   ....G..... .....TTA.. .......... .......... .......... ..........
RG-Y1   ....G..... .....T.A.. .......... .......... .......... ..........
RG-Y2,3 ....G..... .....T.A.. .......... .......... .......... ..........
RG-Y4   ....G..... .....T.A.. .......... .......... .......... ..........

        361                CSB-D                                           420
RM-T1,2 CACGGTGCCC TATTCCTGGC ATTTGGTTCC TATTTCAGGG CCATTACCTG ATATTATTCC
RG-T1   .........A. ...._____. _____..... .......... .......TC. ..-.......
RG-F1   .........A. ...._____. _____..... .......... .......TC. ..-.......
RG-Y1   .........A. ...._____. _____..... .......... .......TC. ..-.......
RG-Y2,3 .........A. ...._____. _____..... .......... .......TC. ..-.......
RG-Y4   .........A. ...._____. _____..... .......... .......TC. ..-.......

        421                                                                480
RM-T1,2 CCCCACTTTC CTTGACCCTG GCATAAG-TT GTTGGTGGAG TACATTTTTA CCCGTGACCC
RG-T1   .........A. .....A.... ....TT.A.. .........G. .T..ACAGG. A.........
RG-F1   .........A. .....A.... ....TT.A.. .........G. .C..ACAGG. A.........
RG-Y1   .........A. .....A.... ....TT.A.. .........G. .C..ACAGG. A.........
RG-Y2,3 .........A. .....A.... ....TT.A.. .........G. .C..ACAGG. A.........
RG-Y4   .........A. .....A.... ....TT.A.. .........G. .C..ACAGG. A.........

        481                                          PY                    540
RM-T1,2 CACATGCCGG GCATTCACTC TAAGGGACAT GGTTATGTTT TTTTTTTTAT TCCTT-CATT
RG-T1   .......... ..G....... ...T...... ....._--___ _____T_. ____T____
RG-F1   .......... ..G....... ...T...... ....._--___ _____T_. ____T____
RG-Y1   .......... ..G....... ...T...... ....._--___ _____GGT. ____T..C.
RG-Y2,3 .......... ..G....... ...T...... ....._--___ _____GGT. ____T..C.
RG-Y4   .......... ..G....... ...T...... ....._--___ _____GGT. ____T..C.
```

FIG. 1. (a).

```
      541                                                                 600
RM-T1,2 GGGCATTTCA GAGTGCAGCG CTAAGGCAAG CTGACAAGGG TGTACATTTC CTTGGTCGCC
RG-T1   T......... C......... .......... .......... A......... ......T.TA
RG-F1   T......... C......... .......... .......... A......... ......T.TA
RG-Y1   T......... C......... .......... .......... A......... ......T.TA
RG-Y2,3 T......... C......... .......... .......... A......... ......T.TA
RG-Y4   T......... C......... .......... .......... A......... ......T.TA

      601                                                                 660
RM-T1,2 GGGCAAAATG TCCATGGTGA AAAGATTTTA TTAGAAGAAC TGAATAACTG ATATCGTGAG
RG-T1   AT........ .T........ .......... .........T ..C....... ..T..A....
RG-F1   AT........ .T........ .......... .........T ..C....... ..T..A....
RG-Y1   AT........ .T........ .......... .........T ..C....... ..T..A....
RG-Y2,3 AT........ .T........ .......... .........T ..C....... ..T..A....
RG-Y4   ATA...... .T........ .......... .........T ..C....... ..T..A....

      661                                                                 720
RM-T1,2 CATAAATAGT TAGTGGTTTC TCCTAAACGA CCTAAGGTGT GG-CCCCTCG GTTTTTGCGG
RG-T1   .......... .......... ......GAT. .........A ..C....C.. ..........
RG-F1   .......... .......... ......GAT. .........A ..C....C.. ..........
RG-Y1   .......... .......... ......GAT. .........A ..C....C.. ..........
RG-Y2,3 .......... .......... ......GAT. .........A ..C....C.. ..........
RG-Y4   .......... .......... ......GAT. .........A ..C....C.. ..........

      721        CSB-2                                    CSB-3           780
RM-T1,2 GTTAAACCCC CCTACCCCCC TAAA-TCGGA AACTATTA-T GATCCTGAAA ACCCCCCGGA
RG-T1   ..A....... .......... ....C...T. .........T ..... ..... ..........
RG-F1   ..A....... .........- ....C...T. .........T ..... ..... ..........
RG-Y1   ..A....... .........- ....C...T. .........T ..... ..... ..........
RG-Y2,3 ..A....... .........- ....C...T. .........T ..... ..... ..........
RG-Y4   ..A....... .........- ....C...T. .........T ..... ..... ..........

      781                                                                 840
RM-T1,2 AACAGGGAAA GTCTCCAGTT AGGAATTTAG C-CTCCCAAA ACGCATCTATT TACATTATT
RG-T1   ....-..... .....G..C. .........-. .T.C..... .......... .........
RG-F1   ....-..... .....G..C. .........-. .T.C..... .......... .........
RG-Y1   ....-..... .....G..C. .........-. .T.C..... .......... .........
RG-Y2,3 ....-..... .....G..C. .........-. .T.C..... .......... .........
RG-Y4   ....-..... .....G..C. .........-. .T.C..... .......... .........

              D-loop   tRNA^Phe
      841       <---++--->                          ***              900
RM-T1,2 AAAATAATAT TTTTGCCAAC GTAGCTTAAC TAAAGCATAA CACTGAAGATG TTAAGATGG
RG-T1   .....G.... .......... .......... .......... .......... .........
RG-F1   .....G.... .......... ..........
RG-Y1   .....G.... .......... ..........
RG-Y2,3 .....G.... .......... ..........
RG-Y4   .....G.... .......... ..........

                        tRNA^Phe    12S rRNA
      901                 <---++--->          940
RM-T1,2 GCCCTAGAAA GCCCCGTAAG CACAAAAGCT TGGTCCTGAC
RG-T1   A......... .T........ .......... ..........
```

FIG. 1. (b).

FIG. 1. Aligned sequences of control region and tRNA^{Phe} of two *Rhinogobius* species. RM, *R. maculafasciatus*; RG, *R. giurinus*. *** the anti-codon of tRNA gene. The underlined region as PY (pyrimidine tracts) and CSB (conserved sequence block).

aligned mtDNA D-loop sequences, compared among eight individuals of the two *Rhinogobius* species, are shown in Fig. 2. The 5′ region (0–300 bp) is more divergent than the following 300 bp. The highest divergence rate (22–34%) occurred in the 51–200 bp of 5′ region. The central region is conserved with only one transition within 315–400 bp; a similarly conserved sequenced block (CSB) has been observed in several teleosts (Lee *et al.*, 1995). The intraspecific divergence within *R. giurinus* is 0·8–1·8% and the 5′ region is more variable than the middle and 3′ regions.

TABLE II. Length and conserved sequence blocks (CSB) of mtDNA D-loop sequences from 19 species of fishes

Order Family	Species	Length	CSB-D	CSB-2	CSB-3
Acipenseriformes					
Acipenseridae	*Acipenser transmontanus*	~976	TTACTGGCA TCTGGTTCC	CAAACCCCC—TACCCCC	TGT–CAAACCCCAA–AAGCA
Clupeiformes					
Clupeidae	*Alosa pseudoharengus*	1017	TTCCTGGCA TTTGGTTCC	CAAACCCCCC–TACCCCC	TGT–AAACCCCGA–AACCA
Salmoniformes					
Salmonidae	*Oncorhynchus mykiss*	1003	TTCCTGGCA TTTGGTTCC	TAAACCCCCC–TACCCCC	TGT–TAAACCCCTA–AACCA
	Salmo salar	1023	TTCCTGGCA TTTGGTTCC	TAAACCCCCC–TACCCCC	TGT–TAAACCCCTA–AACCA
Cyprinoformes					
Cyprinidae	*Cyprinus carpio*	927	TTACTGGCA TCTGGTTCC	CAAACCCCCCTTACCCCC	TGT–CAAACCCCGAAACCAA
Homalopteridae	*Crossostoma lacustre*	896	TTACTGGCA TCTGGTTCC	CAAACCCCCT–TACCCCC	TGCTCAAACCCCGAAACCAA
Gadiformes					
Gadidae	*Gadus morhua*	997	TTCCTGGCT ATTCTG–CC	TAAACCCCCCCTCCCCCC	TGT–AAACCCCCCGGAAACA
	Microgadus tomcod	853	TTCCTGGCT ATTCTG–CC	TAAACCCCCCCCCCCCCC	TTC–CAACCCCCCGGAAACA
	Pollachius virens	868	TTCCTGGCT ATTCTG–CC	TAAACCCCCC–CCCCCCC	TTC–CAACCCCCCGGAAACA
	Melanogrammus aeglefinus	856	TTCCTGGCT ATTCTG–CC	TAAACCCCCC–CCCCCCC	TGC–AAACCCCCCGGGAACA

TABLE II. *Continued*

Order Family	Species	Length	CSB-D	CSB-2	CSB-3
Perciformes					
Gobiidae	*Rhinogobius maculafasciatus*	841	TTCCTGGCA TTTGGTTCC	TAAACCCCCC–TACCCCC	TGA–AAACCCCCCGGAAACA
	Rhinogobius giurinus	841–2	TTCCTGGCA TTTGGTTCC	TAAAACCCCCCTACCCCC	TGA–AAACCCCCCGGAAACA
Cichlidae	*Cichlasoma citrinellum*	888	TTCCTGGCA TTTGGTTCC	TAAACCCCCC–TACCCCC	TGT–AAACCCCCCGGAAACA
Sparidae	*Acanthopagrus australis*	932–3	TTCCTGGCA TTTGGTTCC	TAAACCCCCCCCCCCCCC	TGC–AAACCCCTCAAAAACA
	Sparus sarba	971–3	TTACTGGCA TTTGGTTCC	TAAACCCCCCCCCCCCCC	TGT–AAACCCCTCAGAAACA
Moronidae	*Dicentrarchus labrax*	~2499	ATATTTGCA TTTGGCTGT	TTTGCCCCCCCTACCCCC	TAA–ATCCCCCTAAGAAAAA
Pleuronectiformes					
Pleuronectidae	*Limanda ferruginea*	~1600	TTCCTGGCA TTTGGTTCC	AAAACCCCCC–TACCCCC	TGA–AAACCCCCCGGAAACA
	Hipploglossoides platessoides	1553	TTCCTGGCA TTTGGTTCC	AAAACCCCC–-TACCCCC	TGA–AAACCCCCCGGAAACA
	Pseudopleuronectes americanus	1033	TTCCTGGCA TTTGGTTCC	AAAACCCCC–-TACCCCC	None

TABLE III. The distance matrix of aligned control region sequences among different haplotypes of the *Rhinogobius* samples. The upper half (above diagonal) as mean distance and lower half (below diagonal) as number of divergent sites

	1	2	3	4	5	6
1 RMA-T1,T2	—	0·115	0·113	0·117	0·113	0·114
2 RGU-T1	98	—	0·008	0·016	0·013	0·015
3 RGU-F1	96	7	—	0·018	0·014	0·016
4 RGU-Y1	99	14	15	—	0·004	0·006
5 RGU-Y2,Y3	96	11	12	3	—	0·002
6 RGU-Y4	97	13	14	5	2	—

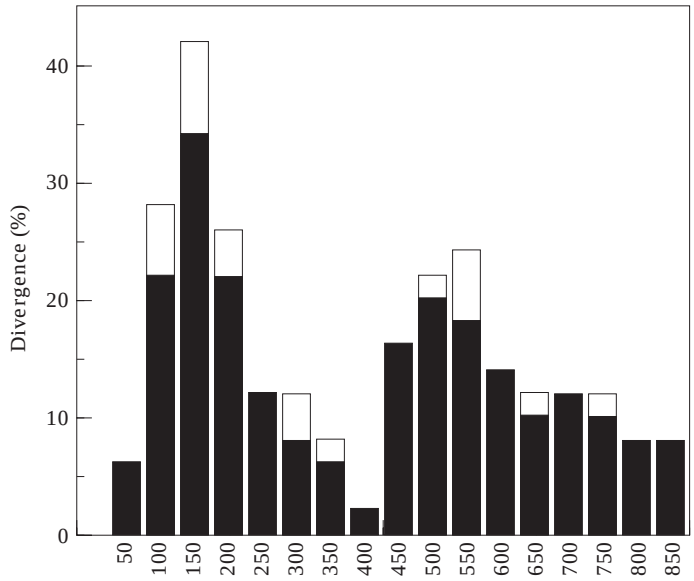

FIG. 2. Percentages of divergence per 50 bp of mtDNA control region compared (a) between two *Rhinogobius* species (■) and (b) the populations within *R. guirinus* (□).

TERMINATION ASSOCIATED SEQUENCES

Termination-associated sequences (TAS) have been observed in many vertebrates as folding sequences which cause termination of strand replication in control regions (Southern, 1988). Two TAS-like sequences were examined in the two *Rhinogobius* species, TAS-1 is located from 6 to 15 bp (ACATATAGT) and the other TAS-2 is located from 71 to 83 bp (ACATCATCTATGT in *R. maculafasciatus* and ACATT(G/A)CTATG in *R. giurinus*) (Fig. 1).

CONSERVED SEQUENCE OF CSB AND PY

Two or three CSBs have been suggested for fish and higher vertebrate sequences (Saccone *et al.*, 1991; Lee *et al.*, 1995). The central regions also contain pyrimidine tracts (PY) in the mammals and fishes (Saccone *et al.*, 1991; Shedlock *et al.*, 1992). In the present gobiid D-loop sequences, the first CSB,

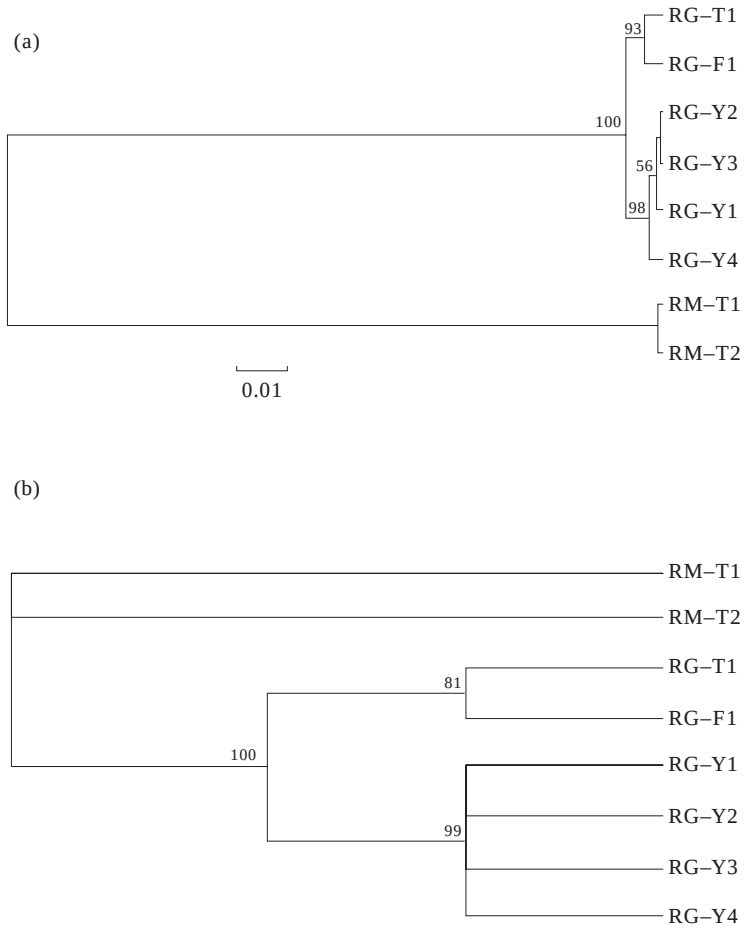

FIG. 3. Phylogenetic cladogram for two *Rhinogobius* species. (a) UPGMA tree by the Kimura two parameter distance and (b) maximum parsimony tree, both trees with value of 1000 bootstrap replications.

CSB-D, is located from the 368 to 385 bp (TTCCTGGCATTTGGTTCC); the second, CSB-2, is located from the 718 (717) to 734 bp (T(A)AAACC CCCCTACCCCC) and a third, CSB-3, is located from the 761 to 779 bp (TGAAAACCCCCCGGAAACA). Gobiid fishes are similar to the Perciformes in the possession of three CSBs (Jean *et al.*, 1995; Lee *et al.*, 1995) (Table II).

MOLECULAR PHYLOGENY

Among eight individuals of rhinogobies, the two *R. maculafasciatus* had an identical haplotype in the D-loop region whereas the six individuals of *R. giurinus* from three localities had five haplotypes (Fig. 1). Both the phylogenetic analyses by distance method (UPGMA) and maximum parsimony supported well the currently recognized two distinct morphological species (Fig. 3). Especially, the level of interspecific differentiation is higher than that of the observed intraspecific variation. There are 96 to 98 mutation sites for comparison between all six haplotypes represented in the eight specimens of both species

but only two to 15 mutation sites among the five different haplotypes within *R. giurinus* (Table III). For intraspecific relationships between the six individuals from the three populations of *R. giurinus*, two significant groups, from Taiwan–Fujian and Yunan, respectively, were indicated in both trees with the same support from high bootstrap values (Fig. 3).

DISCUSSION

The length of mtDNA D-loop sequences varies between different fish taxa (Table 2). The major differences are due to the variable copy numbers of tandemly repeated sequences (Buroker *et al.*, 1990), such as 40-bp repeats in the cod *Gadus morhua* L. (Arnason & Rand, 1992) and 82-bp repeats in the sturgeon *Acipenser transmontanus* Richardson (Buroker *et al.*, 1990). An unusual feature of variability in D-loop length has been found in the European sea bass *Dicentrarchus labrax* (L.) where a wide range in length was observed among individuals, ranging from 853 to 1477 bp, due to great variation in copy numbers of two different units of tandem repeats (Cesaroni *et al.*, 1997), reaching in one individual an extreme length of 2499 bp (Cecconi *et al.*, 1995).

Recently, a compact D-loop sequence of 853 bp has been observed in the tomcod *Microgadus tomcod* (Walbaum) (Lee *et al.*, 1995). The two rhinogoby sequences reported here are slightly shorter at 841–842 bp and lack tandem repeats in the eight fishes examined. These D-loop sequences are the shortest ones of all those published to date for teleost fishes. Recent unpublished work by I.S.C. and P.J.M. indicates similar D-loop lengths of less than 860 bp for other gobioid taxa. An advantage of such compactness in D-loop length among gobies would be to make easier the aligning of sequences for phylogenetic analysis.

Usually three CSBs can be observed in fish D-loop sequences (Table 2) although CSB-3 is lost in *Pseudopleuronectes americanus* (Walbaum) due to disruption by tandem repeats (Lee *et al.*, 1995). These three conserved sequences provide good potential for designing suitable primers for amplifying mtDNA D-loop sequences for phylogenetic use among related fish taxa.

The slowly evolving central domain of the D-loop sequence, which always encloses the CSB-D, seems to offer some opportunity for constructing the phylogeny of higher taxa up to family level (Lee *et al.*, 1995). However, only a few characters among the 168 bp can be used, and phylogenetic results have only low bootstrap support (Lee *et al.*, 1995). It seems more appropriate to use this region for phylogenetic reconstruction below the genus level, even for the detection of intraspecific variation (Meyer *et al.*, 1990; Sang *et al.*, 1994; Yang *et al.*, 1994; Jean *et al.*, 1995) because of likely homoplasy in this rapidly evolving region at higher taxonomic level.

From morphological studies, *Rhinogobius*, as currently defined (Akihito *et al.*, 1993) has been found to contain two distinct species-groups. One of these is the *R. brunneus* group, characterized by a longitudinal pattern of infraorbital sensory papillae and containing at least 40 species including the present *R. maculafasciatus*, and the other is the *R. giurinus* group where infraorbital cheek papillae occur in transverse rows and which comprise merely two species (Masuda *et al.*, 1984; Akihito *et al.*, 1993; Chen & Shao, 1996; Chen & Miller, unpubl.). The *R. brunneus* group has radiated into many distinct morphological

species with a high degree of local endemism (Akihito *et al.*, 1993; Chen & Shao, 1996), in contrast to the *R. giurinus* forms across a similar range (Chu & Wu, 1965; Chen & Shao, 1996; Chen & Miller, unpubl.). Both *R. maculafasciatus* and *R. giurinus* are anadromous species which spawn in fresh water and have a planktonic postlarval stage, which drifts to lower reaches and estuaries. Subsequently, benthic juveniles ascend the rivers to breed (Miyadi *et al.*, 1986; Mizuno & Goto, 1987; Kananabe & Mizuno, 1989; Chen, 1994; Chen & Shao, 1996).

The D-loop sequences from two *Rhinogobius* species examined has shown that they comprise holoplasmic individuals as is the normal condition for those fishes investigated (Lee *et al.*, 1995) and unlike the unusual case of the European sea bass, where there is a high frequency of heteroplasmic individuals and occurrence of D-loop length-polymorphism (Cesaroni *et al.*, 1997). The latter complexities of heteroplasmy and length-polymorphism are disadvantageous for phylogenetic analysis of such species. However, the phenomenon of different copy numbers in tandem repeat does not seem to disturb the function of mtDNA D-loop sequences, and may suggest that length is not so important for influencing D-loop function regardless of the additional copies produced by the tandem repeats of specific parts (Cesaroni *et al.*, 1997). On the other hand, although the compact D-loop length of the present gobioid fishes less than 850 bp, the several conserved sequences present in all fishes can still be observed as mentioned above.

USE OF D-LOOP SEQUENCES IN PHYLOGENETIC RECONSTRUCTION FOR THE GENUS *RHINOGOBIUS*

The compact mtDNA D-loop sequences in rhinogobies are useful genetic markers not only for constructing interspecific phylogeny of recognizably morphological species but also in being sensitive enough for intraspecific phylogenic/phylogeographical reconstruction of population relationships. Overall, the divergence rate of short mtDNA D-loop sequences from the present two gobiids suggests that this region is sufficiently informative for distinguishing cryptic species.

Thus, D-loop sequences determined for *R. maculafasciatus* have features in a molecular biological investigation of species definition and phylogeny within the genus *Rhinogobius*. Results for *R. maculafasciatus* and other members of the *brunneus* species group have shown that this complex is divisible into two main clades, with at least eight valid species from Taiwan and five from Okinawa as well as other species still under investigation from southern mainland China (Chen & Miller, unpubl.).

Regarding intraspecific variation of D-loop sequences in *R. giurinus* populations, the results supported by different methods of sequence comparison suggest that a Yunan group in the upper drainages of the Nanpan River (Pearl River basin) is more isolated than both the Fujian and Taiwan populations. The reason may be that the Yunan population of *R. giurinus*, adapted to the upper reaches of the Nanpan, may have become effectively landlocked and isolated completely from other populations in the main river basins of southern China and even Taiwan, with likely genetic isolation. However, a more complete intraspecific phylogeny of *Rhinogobius* species requires more samples from different localities in China.

We thank the two referees for valuable suggestions concerning this manuscript. One of us (I-S. Chen) thanks the Taiwanese government (Ministry of Education) for support of postgraduate work at Bristol University.

References

Akihito, Iwata, A., Ikeda, Y. & Sakamoto, K. (1993). Suborder gobioidei. In *Fishes of Japan with Pictorial Keys to the Species* (Nakabo, T., ed.), pp. 997–1116. Tokyo: Tokai University (in Japanese).

Anderson, S., Bankwier, A. T., Barrell, M. H. L., Debruijn, A. R., Coulson, J., Drouin, I. C., Eperon, D., Nierlich, P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R. & Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290,** 457–465.

Arnason, E. & Rand, D. M. (1992). Heteroplasmy of short tandem repeats in mitochondrial DNA of Atlantic cod, *Gadus morhua*. *Genetics* **132,** 211–220.

Avise, J. C., Ball, R. M. & Arnold, J. (1988). Current versus historical population size in vertebrate species with high gene flow: a comparison based on mitochondrial DNA lineages and interbreeding theory for neutral mutation. *Molecular Biology and Evolution* **5,** 331–344.

Bibb, M. J., Etten, R. A., Wright, C. T., Walberg, M. W. & Clayton, D. A. (1981). Sequence and gene organization of mouse mitochondrial DNA. *Cell* **26,** 167–180.

Brown, W. M. (1983). Evolution of animal mitochondrial DNAs. In *Evolution of Genes Proteins* (Nei, M. & Koehn, R. K., eds), pp. 62–88. Sunderland, MA: Sinauer.

Brown, W. M., George, M., 6Jr & Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the U.S.A.* **76,** 95–130.

Buroker, N. E., Brown, J. R., Gilbert, T. A., O'Hara, P. J., Beckmanbach, A. T., Thomas, W. K. & Smith, M. J. (1990). Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. *Genetics* **124,** 157–163.

Cecconi, F., Giorgi, M. & Mariottini, P. (1995). Unique features in the mitochondrial D-loop region of the European seabass *Dicentrarchus labrax*. *Gene* **165,** 149–155.

Cesaroni, D., Venanzentti, F., Allegrucci, G. & Sbordoni, V. (1997). Mitochondrial DNA length variation and heteroplasmy in natural population of the European sea bass, *Dicentrarchus labrax*. *Molecular Biology and Evolution* **14,** 560–568.

Chang, Y.-S., Huang, F.-L. & Lo, T.-B. (1994). The complete nucleotide sequence and gene organization of carp (*Cyprinus carpio*) mitochondrial genome. *Journal of Molecular Evolution* **38,** 138–155.

Chen, I-S. (1994). The systematic studies of *Rhinogobius brunneus* complex from Taiwan. Msc. thesis of National Sun Yat-Sen University, Kaohsiung. 114 pp. (in Chinese with English abstract).

Chen, I-S. & Shao, K.-T. (1996). A taxonomic review of the gobiid fish genus *Rhinogobius* Gill, 1859 from Taiwan, with descriptions of three new species. *Zoological Studies* **35,** 200–214.

Chu, X.-L. & Chen, Y.-R. (eds) (1990). *The Fishes of Yunan, China.* Vol. II. Beijing: Science Press (in Chinese).

Devereux, J., Haeberli, P. & Marquess, P. (1991). *Genetic Computer Group Manual, Revision 7.0.* Madison, WI: University of Wisconsin.

Digby, T. J., Gray, M. W. & Lazier, C. B. (1992). Rainbow trout mitochondrial DNA: sequence and structure characteristics of the non-coding control region and flanking tRNA genes. *Gene* **113,** 197–204.

Felsenstein, J. (1993). *PHYLIP: Phylogenetic Inference Package, version 3.5.* Seattle: University of WA.

Gill, T. N. (1859). Notes on a collection of Japanese fishes by Dr J. Morrow. *Proceedings of the Academy of Natural Sciences of Philadelphia* **1859,** 144–149.

Iguchi, K. & Mizuno, N. (1990). Diel changes of larval drift among amphidromous gobies in Japan, especially *Rhinogobius brunneus*. *Journal of Fish Biology* **37,** 255–264.

Jean, C.-T., Hui, C.-F., Lee, S.-C. & Chen, C.-T. (1995). Variation in mitochondrial DNA and phylogenetic relationships of fishes of the sudfamily Sparinae (Perciformes: Sparidae) in the coastal waters of Taiwan. *Zoological Studies* **34,** 270–280.

Johansen, S., Guddal, P. H. & Johanson, T. (1990). Organization of the mitochondrial genome of Atlantic cod, *Gadus morhua*. *Nucleic Acids Research* **18,** 411–419.

Kawanabe, H. & Mizuno, N. (1996). *Freshwater Fishes of Japan.* Tokyo: Yama-kei Publishers (in Japanese).

Lee, W.-J., Conroy, J., Hunting Howell, W. & Kocher, T. D. (1995). Structure and evolution of teleost mitochondrial control region. *Journal of Molecular Evolution* **41,** 54–66.

Meyer, A., Kocher, T. D., Basabibwaki, P. & Wilson, A. C. (1990). Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial sequences. *Nature* **347,** 550–553.

Miyadi, D., Kawanabe, H. & Mizuno, N. (1986). *Colored Illustrations of the Freshwater Fishes of Japan.* Osaka: Hoikusha Press (in Japanese).

Mizuno, N. (1960). Description of a new freshwater gobies in Japan. *Memoirs of the College of Sciences, University of Kyoto, Series B* **27,** 97–119.

Mizuno, N. & Goto, A. (1987). *Japanese Freshwater Fishes, and Their Distribution, Variation and Speciation.* Tokyo: Tokai University (in Japanese).

Mizuoka, S. (1974). Studies on the variation of *Rhinogobius brunneus* (Temminck et Schlegel) III. On the six types of body color and color pattern in the Sanin district, the Hokuriku district, Sanyo district, and the Goto Islands. *Bulletin of the Faculty of Education, Hiroshima University* **23,** 31–40 (in Japanese with English summary.

Rutter, C. (1987). A collection of fishes obtained in Swatow, China, by Miss Adele M. Fielde. *Proceedings of the Academy of Natural Sciences of Philadelphia* **1897,** 56–90.

Saccone, C., Pesole, G. & Sbisa, E. (1991). The main regulatory region of mammalian mitochondrial DNA: structure–function model and evolutionary pattern. *Journal of Molecular Evolution* **33,** 83–91.

Sambrook, J., Fritsch, E. F. & Manitais, T. (1989). *Molecular Cloning. A Laboratory Manual.* 2nd edn. Cold Spring Harbor: Cold Spring Harbor Laboratory.

Sang, T.-K., Chang, H.-Y., Chen, C.-T. & Hui, C.-F. (1994). Population structure of Japanese eel, *Anguilla japonica*. *Molecular Biology and Evolution* **11,** 250–260.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the U.S.A.* **74,** 5463–5467.

Shedlock, A.-M., Parker, J.-D., Crispin, D. A., Pietsch, T. W. & Burner, T. W. (1992). Evolution of the salmonid mitochondrial control region. *Molecular Phylogenetics and Evolution* **1,** 179–192.

Southern, S., Peter, J. S. & Andrew, E. D. (1988). Molecular characterization of cloned mitochondrial genome. *Journal of Molecular Evolution* **28,** 32–42.

Swofford, D. L. (1993). *PAUP: Phylogenetic Analysis Using Parsimony*. Version 3.1. Champaign, IL: Illinois Natural History Survey.

Tzeng, C.-S., Hui, C.-F., Shen, S.-C. & Huang, P.-C. (1992). The complete nucleotide sequence of *Crossostoma lacustre* mitochondrial genome: conservation and variation among vertebrates. *Nucleic Acids Research* **20,** 4853–4858.

Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U. B., Helm-Bychowski, K. M., Higuchi, R. G., Palimbi, S. R., Prager, E. M., Sage, R. D. & Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of Linnean Society* **26,** 375–400.

Yang, Y.-J., Lin, Y.-S., Wu, J.-L. & Hui, C.-F. (1994). Variation in mitochondrial DNA and population structure of the Taipei treefrog *Rhacophorus taipeianus* in Taiwan. *Molecular Ecology* **3,** 219–228.