

THE MOLECULAR PHYLOGEOGRAPHY OF *CANDIDIA BARBATA* SPECIES COMPLEX (TELEOSTEI: CYPRINIDAE) FROM TAIWAN

Jui-Hsien Wu, Chi-Hsin Hsu

Institute of Marine Resources, National Sun Yat-sen University, Kaohsiung 804, Taiwan, ROC.

Lee-Shing Fang

*National Museum of Marine Biology & Aquarium, Pingtung 944, Taiwan, ROC;
and Cheng-Shiu University, No. 840 Cheng-Ching Rd., Niau-Song, Kaohsiung County 833, Taiwan, ROC.*

I-Shiung Chen

Institute of Marine Biology, National Taiwan Ocean University, Keelung 202, Taiwan, ROC.

Email: isc@mail.ntou.edu.tw

(Corresponding Author)

ABSTRACT. – The species *Candidia barbata*, collected from different river basins in Taiwan, was surveyed for mitochondrial genetic haplotypes. The species is endemic to Taiwan and is widely-distributed around the island. A total of 208 samples of *C. barbata* specimens yielded 53 mitochondrial haplotypes of D-loop sequences. Using congruent tree topology, two major clades were detected from Taiwan: Clade A (North and West) and Clade B (South). The results were obtained by employing both distance (neighbor-joining) and discrete (maximum parsimony) methods of sequence analysis. The mitogenetic diversity is high in this species complex.

KEY WORDS. – *Candidia barbata*, Cyprinidae, species complex, molecular evolution, biogeography, Taiwan.

INTRODUCTION

The small-sized cyprinid genus *Candidia* (Jordan & Richardson, 1909) has been regarded as endemic to Taiwan (Chen, 1982; Tseng, 1986; Chen & Fang, 1999). However, it shares close resemblance with the Japanese *Zacco* taxon. The species *Candidia barbata* (Regan, 1908) was originally described as *Opsariichthys barbatus*. However, the species was validated in the genus *Candidia* by Chen (1998). Chen & Fang (1999) listed the distribution pattern of *Candidia* in Taiwan and it can be found from Taipei County to Pingtung County, mainly in the Western drainages next to the Central Mountain Ridge of Taiwan. This species inhabits small tributaries of river basins (Chen & Fang, 1999; Chen & Chang, 2005).

Taiwan, the largest island off the mainland of China, was formed and lifted by the interaction of the Euro-Asia and Philippine Sea Plates. The waters are shallow on the Formosan bank on the West and very deep (up to 1,000 – 2,000 m) on the Eastern slope. The freshwater river systems mainly flow to the Western and Eastern coasts, being well separated by the Central Mountain Ridge (Page & Suppe, 1981).

The Order Cypriniformes, being primarily freshwater fishes, are strongly affected by the glaciations and river connection via land bridges into the various river basins between mainland China and the island of Taiwan. Low temperatures affect the inland-water fish communities, causing the communities to migrate to the lower reaches from the mainland river systems to avoid temperature stress. The communities then connect to nearby tributaries of other rivers which were once joined as ancient river basins (Froufe et al., 2002; Kotlik & Berrebi, 2002; Perdices et al., 2003).

The mitochondrial DNA (mtDNA) exhibits a high evolutionary rate with the accumulation of neutral mutations and thus, accounts for evolutionary events. Herein, we use the D-loop sequence (or control region) as the DNA genetic marker. This sequence has an AT-rich non-coding region in most species. Since the sensitivity of such DNA markers is good enough to detect the polymorphic haplotypes among species, it is possible to reconstruct the distribution pattern of wide-ranging species by using more detailed molecular phylogeographic and morphological studies (Chen et al., 1998, 2002). This paper reports on the genetic diversity and geographical speciation patterns of the *C. barbata* species complex in Taiwan and hypothesizes as to whether the

speciation event of this species complex is based on monophyletic species or resulting from the isolation of sibling species.

MATERIALS AND METHODS

A total of 208 individuals of *Candidia barbata* were collected from 29 river basins in Taiwan (Fig. 1 & Table 1). Two morphologically similar species, *Zacco temminckii* (Temminck & Schlegel, 1846) and *Zacco sieboldii* (Temminck & Schlegel, 1846), which have been recently revalidated by Hosoya et al. (2003), were assigned as comparative outgroup taxa.

All DNA samples were extracted using the general protocols of the phenol-chloroform method (Sambrook et al., 1989; Chen et al., 2002). The DNA fragments of about 1,500 base pairs (bp), inclusive of the full length D-loop region were amplified by polymerase chain reaction (PCR) using a set of primers for the flanking regions which were designed from the sequences of tRNA-PHE and 12S rRNA (Chen et al.,

1998, 2002). PCR was done in a Model 9700 thermal cycler (Perkin-Elmer) with 30 - 40 cycles carried out for each 25 μ L reaction volume which contained 14.4 μ L sterile distilled water, 2.5 μ L 10 $\times$ PCR buffer (Takara), 2.0 μ L dNTP (2.5 μ M each), 2.5 μ L each primer (5 μ M each), 0.1 μ L 0.5 units Ex *Taq* (Takara) and 1.0 μ L template DNA. Each thermal cycle profile was as follows: denaturation at 94°C for 15 seconds, annealing at 50°C for 15 seconds and extension at 72°C for 60 seconds. The PCR products were electrophoresed on a 1.0% agarose gel (Takara) and which was then stained with ethidium bromide for band characterization via ultraviolet trans-illumination.

The double-stranded PCR products were purified using a kit (Boehringer Mannheim, High Pure PCR Product Purification kit) and were subsequently used for direct cycle sequencing with fluorescent dye-labeled terminators (ABI Big-Dye kit). The primers used were the same as those used for PCR. All sequencing reactions were performed according to the manufacturers' instructions. Labeled fragments were analyzed using an ABI Model 377-64 DNA sequencer (ABI).

Distinction of divergent haplotypes of all nucleotide sequences for *Candidia* species were conducted using Molecular Evolutionary Genetics Analysis program (MEGA) version 3.0 (Kumar et al., 2004). Evolution model testing was conducted using Modeltest version 3.7 (Posada & Crandall, 1998). Phylogenetic and molecular evolutionary analyses were performed using PAUP version 4.0b10 (Swofford, 2002).

The minimum-spanning network was constructed with the MINSPNET (Excoffier & Smouse, 1994). Inter- and intra-population genetic diversity were quantified by indices of haplotype diversity (Nei & Tajima, 1983) and estimates of nucleotide diversity (Jukes & Cantor, 1969) using DnaSP version 3.14 (Rozas & Rozas, 1999).

RESULTS

A total of 208 *Candidia barbata* samples from 29 different river basins in Taiwan were analyzed with two outgroups (*Zacco temminckii* and *Z. sieboldii*). There were 53 distinct haplotypes from the mtDNA D-loop sequences which could be divided into two distinct clades. Clade A comprises of 42 haplotypes (GenBank accession numbers: DQ489379 to DQ489420) and Clade B comprises of 11 haplotypes (GenBank accession numbers: DQ489421 to DQ489431).

The length of complete D-loop sequences is 995 - 997 bp in Clade A, 996 bp in Clade B, 996 bp in *Zacco temminckii* and 997 bp in *Z. sieboldii*. The aligned mtDNA sequences for all haplotypes were up to 1,000 bp. The most common haplotypes were as follows: A11, A30 (both based on 32 individuals), A26 (based on 25 individuals) and A40 (based on 16 individuals). The other haplotypes each occurred in less than 10 individuals (Table 1). Among the distribution patterns of haplotypes, the most widely-distributed was A11 (found in seven river basins), followed by A30 (found in four

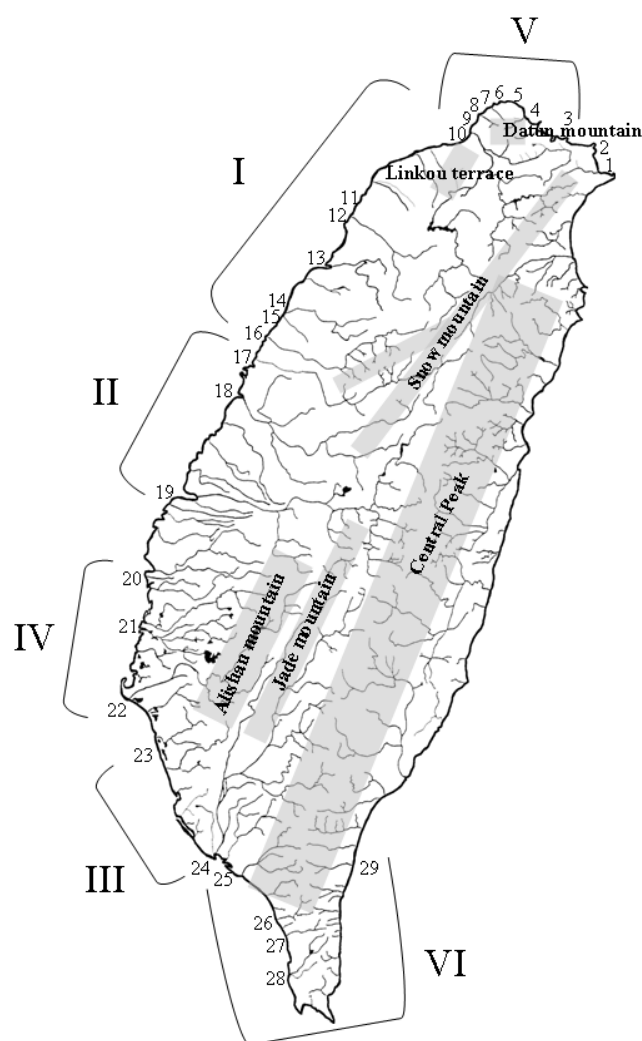


Fig. 1. Sampling sites of *Candidia barbata* for mitogenetic D-loop sequences analysis, showing Grades I to VI. The gray regions depict mountainous areas that might be potential geographical barriers. Numbers represent rivers and names are indicated in Table 1.

Table 1. Distribution of haplotypes from mitochondrial D-loop sequence of *Candidia barbata* complex in Taiwan.

River Basin	No. of Samples	Haplotype	Clade																							
			A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
1 Shuanshi	5	2	2	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2 Beisukun	6	1	-	-	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3 Kulinkun	6	2	-	-	-	4	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4 Yuantan	3	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5 Lawmei	1	1	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6 Tatan	1	1	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7 Shinghuaten	7	3	1	-	-	-	-	-	3	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8 Gueroushan	6	3	-	-	-	-	-	-	-	4	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9 Konzuten	1	1	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10 Tanshuei	17	9	-	-	-	-	-	-	-	-	1	6	2	1	1	3	1	1	1	-	-	-	-	-	-	-
11 Fongshan	7	2	-	-	-	-	-	-	-	-	-	6	-	-	-	-	-	-	1	-	-	-	-	-	-	-
12 Tochien	7	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5	1	1	-	-	-	-	-
13 Chonkang	8	2	-	-	-	-	-	-	-	-	-	7	-	-	-	-	-	-	-	-	-	-	-	1	-	-
14 Holong	3	2	-	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	1	-
15 Sihuh	5	1	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16 Taan	10	1	-	-	-	-	-	-	-	-	-	10	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Total			3	3	6	4	2	1	1	1	3	3	7	32	8	2	1	1	3	1	1	1	5	1	1	1

Table 1. (Continued)

River Basin	No. of Samples	Haplotype	Clade																											
			A26	A27	A28	A29	A30	A31	A32	A33	A34	A35	A36	A37	A38	A39	A40	A41	A42	B01	B02	B03	B04	B05	B06	B07	B08	B09	B10	B11
17 Tajar	6	1	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	B	B	B	B	B	B	B
18 Wu	17	3	6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
19 Joshui	6	2	14	2	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
20 Parchan	11	3	5	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
21 Gishui	3	1	-	-	-	-	7	1	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
22 Tzengwen	21	5	-	-	-	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
23 Erren	8	4	-	-	-	-	17	-	-	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
24 Kaoping	23	4	-	-	-	-	-	-	-	-	-	-	1	2	1	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-
25 Tongkan	3	3	-	-	-	-	5	-	-	-	-	-	-	-	-	16	1	1	-	-	-	-	-	-	-	-	-	-	-	-
26 Fansan	3	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	1	1	-	-	-	-	-	-	-	-	-
27 Fongkan	5	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	1	-	-	-	-
28 Szuchung	4	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	1	1	-	-
29 Taju	5	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	-	-	-	-	-	1
Total	208	53	25	2	1	1	32	1	3	1	1	1	2	2	1	4	16	1	1	1	1	1	1	9	1	2	1	1	1	1

river basins), A12 and B06 (each found in three river basins) and the rest were found in less than two river basins.

The phylogenetic analysis, reconstructed from the different haplotypes, was conducted using the distance method [neighbor-joining (NJ) method based on the GTR model which was chosen by Modeltest from the sequence data matrix] and the discrete method [maximum parsimony (MP) method by gap complete deletion with 989 bp]. Both methods yielded very similar topology and supported the same six grades (Clade A: Grade I, II, III, IV, V; Clade B: Grade VI) (Fig. 2). The MP tree had 136 trees with equal minimal lengths of 279. The majority-rule consensus shows the consensus tree with NJ tree topology with 1,000 bootstrap replications (CI = 0.62, RI = 0.91). The minimum spanning network was also reconstructed by the six distinct grades and a long branch for Clade B (Fig. 3). From the estimation for divergence times of mitochondrial D-loop sequences by Johns & Avise (1998), this was 5% per million years. Cortey & Garcia-Marin (2002) estimated that the Atlantic brown trout divergence rate was 2%. Wang et al. (2004) used the molecular clock for the cyprinid, *Varicorhinus barbatulus* (= *Onychostoma barbatulus*) at around 3.2 - 10.4%.

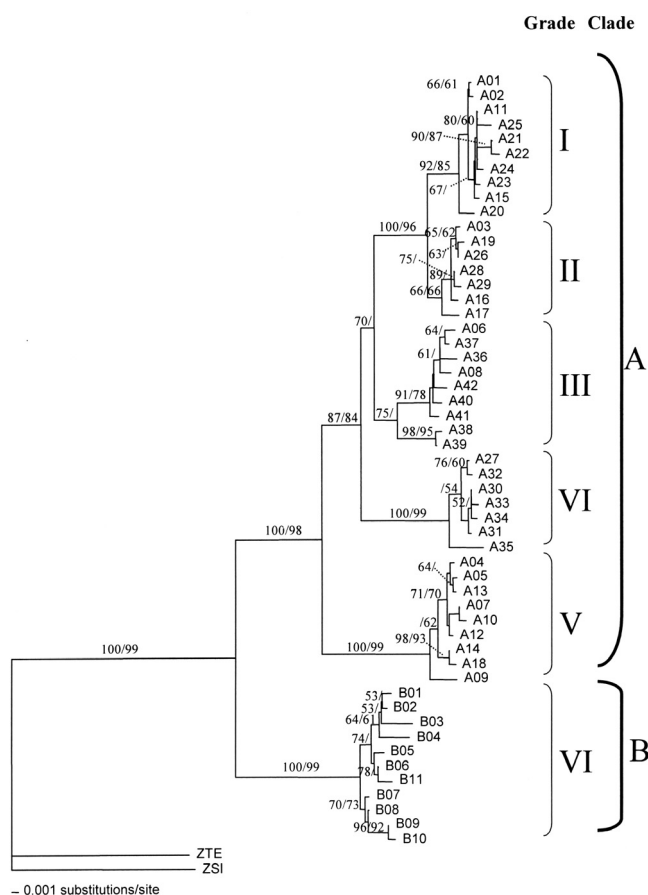


Fig. 2. Neighbor-joining tree of mtDNA haplotypes based on sequence variation of the D-loop region in *Candidia barbata* showing the divisions into six grades (I - VI). Numbers at the nodes are the bootstrap percentage values for 1,000 replicates by neighbour-joining (NJ) and maximum parsimony (MP) methods. *Zacco temminckii* (ZTE) and *Z. sieboldii* (ZSI) are used as the outgroups.

This study employed the estimation as an average of genetic divergence of 6% per million years. Hence, the separation of Clades A and B can be estimated to have occurred 450,000 years ago and the separation between all the *C. barbata* populations and the assigned outgroup, the *Z. temminckii* complex, was 750,000 years ago. All of the events happened after late Pleistocene.

DISCUSSIONS

The *Candidia barbata* species complex in Taiwan is more similar to the Japanese *Zacco temminckii* and *Z. sieboldii* in overall morphology and genetic homology than to other members of *Zacco* (Hosoya et al., 2003). These groups share a similar preference of habitat, inhabiting the upper reaches of rivers and small tributaries rather than the main river basins. However, congeneric or sibling species of *Candidia* have yet to be found in mainland China. From our unpublished data based on recent surveys, *Z. temminckii* may be considered the sister group of the Taiwanese populations of *Candidia*. The generic position of *Z. temminckii* is still in need of clarification. More detailed morphological and osteological studies will be needed to decide if *Z. temminckii* is to be placed in a genus separate from *Zacco* or reassigned as a *Candidia* species.

The six-grade division was generally congruent with the geographic pattern of Taiwan. Clade B (Grade VI) in Hanchuan Peninsula is well-isolated from the other

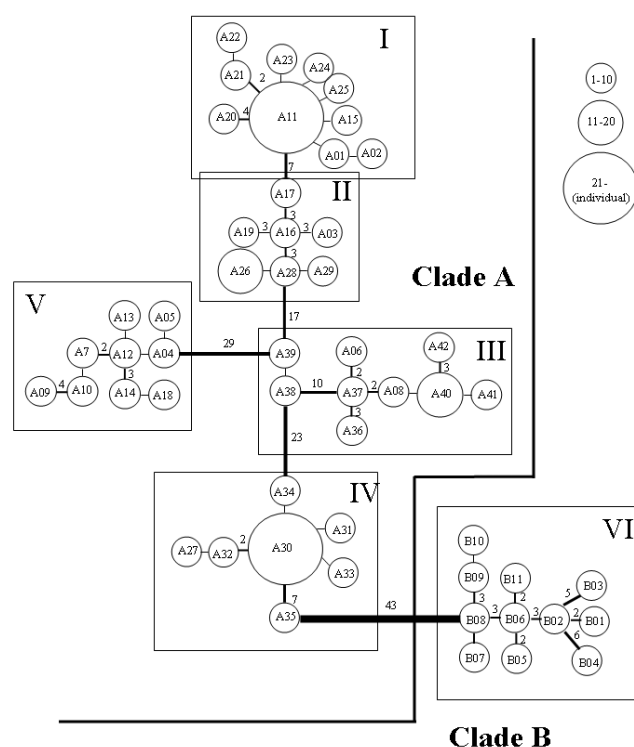


Fig. 3. The minimum-spanning network made by mtDNA haplotypes of the *Candidia barbata* complex. I - VI = grades; haplotype (small circle) = 1 - 10 individuals; haplotype (medium circle) = 11 - 20 individuals; haplotype (large circle) = more than 20 individuals.

geographical regions. Grade V is limited to Northern Taiwan, which is well-separated from the other grades (I, II, III and IV) in Clade A. Among them, Grades I and II are much closer to each other. Grade IV is within the Southern region of the Alisan Ridge which is far from Grade III in the Northern region of the Alisan Ridge. This separation of the Southern Clade B can be explained by the long-term geographical isolation and leading to a greater differentiation between such populations, which may be considered as discrete allopatric species.

Analysis of the sediment Oxygen-18 isotope (^{18}O) records around the South China Sea have revealed the occurrence of 20 glacial periods (marine isotope stages, MIS) within the past 800,000 years (Boulay et al., 2003). From the mitogenetic divergence rates of *C. barbata*, the earliest separation between the Taiwanese *C. barbata* complex and the Japanese *Z. temminckii* and *Z. sieboldii* can be estimated to be around 750,000 years (during the MIS 17 period). The separation between Clade A and B was around 450,000 years ago (at the start of MIS 11), which can be considered the glacial period during MIS 12 and allowed connection by land between mainland China and Taiwan by sea level regression. This land connection might be the isolating mechanism for Clades A and B. After this glacial period, the coastline of the Southern region of Taiwan seems to be well-isolated between these geographical regions during the inter-glacial period and all the events of the glacial periods that follow.

The species complex of two distinct clades was well supported by both distance and discrete methods. The mitogenetic diversity is still very high suggesting a good genetic pool still exists. The different geographical regions seem to have independently-evolved mitogenetic haplotypes as Clade B was well-differentiated from Clade A from the molecular sequence evidence. The allopatric group, *Candidia* sp. in the South, had been recently examined in detail and exhibits distinct morphological and meristic features and is currently being described (Chen et al., unpublished data). The remaining main group, inclusive of the North and West groups, would thus be assigned as *C. barbata*. The occurrence of this endemic species complex in Taiwan serves as an important example when conservation of freshwater fish communities should be considered.

ACKNOWLEDGEMENTS

We are very grateful to Dr. T. Y. Chiang and Mr. C. H. Yang for their kind assistance in the use and discussions of the DNA analysis software. This study is supported by a grant for ISC from National Science Council, Taipei, ROC. This study is also supported by Institute of Zoology, Academia Sinica, Taipei.

LITERATURE CITED

- Boulay, S., C. Colin, A. Trentesaux, F. Pluquet, J. Bertaux, D. Blamart, C. Buehring & P. Wang, 2003. Mineralogy and sedimentology of Pleistocene sediment in the South China Sea (ODP Site 1144). In: Prell, W. L., P. Wang, P. Blum, D. K. Rea & S. C. Clemens, (eds.), *Proceeding of the Ocean Drilling Program, Scientific Results*, **184**: 1-21.
- Chen, I-S. & Y. C. Chang, 2005. *The photographic guide of inland water fishes. Vol. I. Cypriniformes*. Sheichuan Press, Keelung. 284 pp.
- Chen, I-S. & L. S. Fang, 1999. *The Freshwater and Estuarine Fishes of Taiwan*. National Museum of Marine Biology & Aquarium Press, Pingtung. 287 pp. (In Chinese).
- Chen, I-S., C. H. Hsu, C. F. Hui, K. T. Shao, P. J. Miller & L. S. Fang, 1998. Sequence length and variation of mitochondrial control region from two freshwater gobiid fishes belonging to *Rhinogobius* (Teleostei: Gobiidae). *Journal of Fish Biology*, **53**: 179-191.
- Chen, I-S., P. J. Miller, H. L. Wu & L. S. Fang, 2002. Taxonomy and mitochondrial sequence evolution in non-diadromous species of *Rhinogobius* (Teleostei: Gobiidae) of Hainan Island, southern China. *Marine & Freshwater Research*, **53**: 259-273.
- Chen, Y., 1982. A revision of opsariichthine cyprinid fishes. *Oceanologia et Limnologia Sinica*, **13**(3): 293-299. (In Chinese).
- Chen, Y., (ed.), 1998. *Fauna Sinica. Cypriniformes Vol. II*. Science & Technology Publishers, Beijing. 531 pp. (In Chinese).
- Cortey, M. & J.-L. Garcia-Marin, 2002. Evidence for phylogeographically informative sequence variation in the mitochondrial control region of Atlantic brown trout. *Journal of Fish Biology*, **60**: 1058-1063.
- Excoffier, L. & P. E. Smouse, 1994. Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. *Genetics*, **136**: 343-359.
- Froufe, E., I. Magyary, I. Lehoczy & S. Weiss, 2002. mtDNA sequence data supports an Asian ancestry and single introduction of the common carp into the Danube Basin. *Journal of Fish Biology*, **61**: 301-304.
- Hosoya, K., H. Ashiwa, M. Watanabe, K. Mizuguchi & T. Okazaki, 2003. *Zacco sieboldii*, a species distinct from *Zacco temminckii* (Cyprinidae). *Ichthyological Research*, **50**: 1-8.
- Johns, G. C. & J. C. Avise, 1998. Test for ancient species flocks based on molecular phylogenetic appraisals of *Sebastes* rockfishes and other marine fishes. *Evolution*, **52**: 1135-1146.
- Jordan, D. S. & R. E. Richardson, 1909. A catalogue of the fishes of the islands of Formosa, or Taiwan, based on the collections of Dr. Hans Sauter. *Memoirs Carnegie Museum*, **4**(4): 159-204.
- Jukes, T. H. & C. R. Cantor, 1969. Evolution of protein molecules. In: Munro, H. N. (ed.), *Mammalian Protein Metabolism, III*. Academic Press, New York. Pp. 21-132.
- Kotlik, P. & P. Berrebi, 2002. Genetic subdivision and biogeography of the Danubian rheophilic barb *Barbus petenyi* inferred from phylogenetic analysis of mitochondrial DNA variation. *Molecular Phylogenetics and Evolution*, **24**: 10-18.
- Kumar, S., K. Tamura & M. Nei, 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Briefings in Bioinformatics*, **5**: 150-163.
- Nei, M. & F. Tajima, 1983. Maximum likelihood estimation of the number of nucleotide substitutions from restriction sites data. *Genetics*, **105**: 207-217.

- Page, B. M. & J. Suppe, 1981. The Pliocene Lichi melange of Taiwan: Its plate-tectonic and olistostromal origin. *American Journal of Science*, **281**: 193-227.
- Perdices, A., I. Doadrio, P. S. Economidis, J. Bohlen & P. Banarescu, 2003. Pleistocene effect on the European freshwater fish fauna: double origin of the cobitid genus *Sabanejewia* in the Danube basin (Osteichthyes: Cobitidae). *Molecular Phylogenetics and Evolution*, **26**: 289-299.
- Posada, D. & K. A. Crandall, 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**(9): 817-818.
- Regan, C. T., 1908. Description of new species from Lake Candidia, Formosa, collected by Dr. A. Moltrecht. *Annals and Magazine of Natural History*, (Series 8), **2**(10): 358-360.
- Rozas, J. & R. Rozas, 1999. DnaSP version 3.0: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**: 174-175.
- Sambrook, J., E. F. Fritsch & T. Maniatis, 1989. *Molecular Cloning*. 2nd edition. Cold Spring Harbor Laboratory Press, New York. 1659 pp.
- Swofford, D. L., 2002. PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sinauer Associates, Sunderland, Massachusetts.
- Temminck, C. J. & H. Schelegel, 1846. *Pisces in Siebold's Fauna Japonica*. Lunduni Batavorum, Leiden. 323 pp, 143 pls.
- Tseng, C. S., 1986. *Freshwater Fishes of Taiwan*. Taiwan Education Ministry Press, Taichung. 197 pp. (In Chinese).
- Wang, J. P., H. D. Lin, S. Huang, C. H. Pan, X. L. Chen & T. Y. Chiang, 2004. Phylogeography of *Varicorhinus barbatulus* (Cyprinidae) in Taiwan based on nucleotide variation of mtDNA and allozymes. *Molecular Phylogenetics and Evolution*, **31**: 1143-1156.