Genetic Diversity of the Vulnerable *Pangasius sanitwongsei* using Microsatellite DNA and 16S rRNA

Uthairat Na-Nakorn[1], Srijanya Sukmanomon[2], Supawadee Poompuang[1],Panya Saelim[1], Nobuhiko Taniguchi[3] and Masanishi Nakajima[3]

ABSTRACT

Sixty-six *Pangasius sanitwongsei* were collected from five sites, namely Chiangkhong District, Chiangrai Province (CR; N = 1, for mt-DNA study only), Muang District, Nakornphanom Province (NP-1; N = 4), Tatphanom District, Nakornphanom Province (NP-2; N = 41), Khongjiam District, Ubonratchathani Province (UR; N = 10) and Koagsrisuphan District, Sakonnakorn Province (SN; N = 10). Seven microsatellite primers developed from *Pangasianodon gigas* were used. The results indicated very low genetic variation of *P. sanitwongsei* (A = 1.43 - 3.0; H_O = 0.11 - 0.30; H_e = 0.12 - 0.36). Although the least diverse population (SN, A = 1.43; H_O= 0.11; H_e = 0.12) might be of hatchery origin, the genetic variation of the other populations was still low. Population structuring was significant (F_{ST} = 0.190). The genotypic differentiation was significant between SN and the others. Genetic distance was between 0.0952 - 0.4872. The sequences of the 571 bp16S rRNA gene region revealed a total of 6 haplotypes (CR = 1; NP-1 = 1; NP-2 = 5; SN = 2; UR = 1) with 8 variable sites comprising 5:3 transitional to transversional mutation. Haplotype and nucleotide diversity varied from 0 (NP-1 and UR) to 0.3556 (SN) and 0 (CR, NP-1, UR) to 0.0006 (NP-2, SN), respectively. The status of *P. sanitwongsei* as an endangered species was validated. At least two evolutionary significant units (ESU) were identified i.e. NP-1+NP-2 and UR. Samples of SN were mainly of hatchery origin and closely related to CR as evidenced by 16S rRNA.

INTRODUCTION

There are at least 11 of the total 21 species of Pangasiid catfishes which are distributed in Southeast Asia (Roberts and Vidthayanon, 1991). Among these is *Pangasius sanitwongsei*, which can grow up to 300 kg (Roberts and Vidthayanon, 1991), and second in size to *Pangasianodon gigas*. Its distribution expands across the Mekong and Chaophraya Rivers whereby it is either no longer common (Roberts and Vidthayanon, 1991), or, even extinct in the latter. Hence it has been listed as an endangered species in Thailand (OEPP, 1997), although labeled as data deficient (1 August 1996) by the

[1] Department of Aquaculture, Faculty of Fisheries, Kasetsart University, Bangkok, Thailand
[2] Department of Fisheries, Ministry of Agriculture and Cooperatives, Bangkok, Thailand
[3] Department of Applied Population Genetic Bioinformatics, Tohoku University, Sendai, Japan

International Union for Conservation of Nature (IUCN) (Froese and Pauly, 2005). In addition to its importance in the ecosystem it is potentially useful for aquaculture due to its fast growth (Roberts and Vidthayanon, 1991).

Mekong River is under threat due to human activities for development (Mitchell and Braun, 2003), such as dam construction, fishing and habitat destruction. The decline of fish fauna in this river due to the above activities is well documented (Jensen, 1996). The threats are more likely to be serious for large fishes such as *P. gigas* and *P. sanitwongsei* which need a great stretch of the river for seasonal migration (Poulsen, 2005) and specific ecological niches to complete their life cycle. Moreover, due to the long maturation period, a number of stock might be harvested before they could reproduce. Therefore it is certain that *P. sanitwongsei*, which has already disappeared from the Chaophraya River, is also facing the threat of extinction in the Mekong River.

There are emerging conservation activities for fishes in the Mekong River, hence studies on genetic diversity of fish species are required to identify evolutionary significant units (ESU), and to understand the current status of their genetic variations. Information on genetic diversity of *P. sanitwongsei* is limited, and only the monomorphism of 25 allozyme loci in 3 individuals has been reported by Pouyaud et al. (2000).

This study was conducted using seven microsatellite loci and sequences of 571 bp of the 16S rRNA as markers to verify diversity of 66 *P. sanitwongsei* in Thailand.

MATERIALS AND METHODS

Sample collection

Due to its rarity and migratory behavior there was difficulty in collecting adult fish directly from the Mekong River. Therefore samples from cages and ponds stocked with fingerlings of mixed species of *Pangasius* from the River were the main sources of our samples. Forty-one fish were obtained from 3 cages in Tatphanom District, Nakornphanom Province (NP-2) and 10 fish from Sakonnakorn Fisheries Station, Sakonnakorn Province (SK). A total of 15 adult fish were directly collected from the Mekong, 4 fish from Tatphanom District but at a site 20 km away from the first site (NP-1), 10 fish from Kongjiam District, Ubon Ratchathani Province (UR), and one fish collected from Chiangkhong District, Chiangrai Province (CR) (Table 1). The collection sites are shown in figure 1.

A small piece of the caudal fin was cut and preserved in 90% ethanol and delivered to the Fish Genetics Laboratory, Department of Aquaculture, Faculty of Fisheries, Kasetsart University. Before the fin tissue was taken the fish were identified following Roberts and Vidthayanon (1991) and Smith (1945).

Table 1. Details for collection sites of *Pangasius sanitwongsei* used for this study

Pop. Abbrev.	Collection sites	Lat./Long.	sample size (n)	Description of samples
CR	Chiangkhong District, Chiangrai province	Lat.20º 20' N Long. 100º 50' E	1*	Adult, from Mekong River
NP-1	Muang District, Nakornphanom Province	Lat. 17º 4' 48" N Long. 105º E	4	Adult, from Mekong River
NP-2	Tatphanom District, Nakornphanom Province	Lat. 16º 51' 42" N Long. 104º 48' E	41	Adult from 3 cages stocked with fingerlings collected from Mekong River
UR	Khongjiam District, Ubon Ratchthani Province	Lat. 15º 26' 54" N Long. 105º 30' 28" E	10	Adult, from Mekong River
SN	Koagsrisuphan District, Sakonnakorn Provinc	Lat. 17º 4' 37" N Long. 104º 4' E	10	Adult from a pond stocked with fingerlings from Mekong River and some F1

Note: * used for the 16S rRNA study only

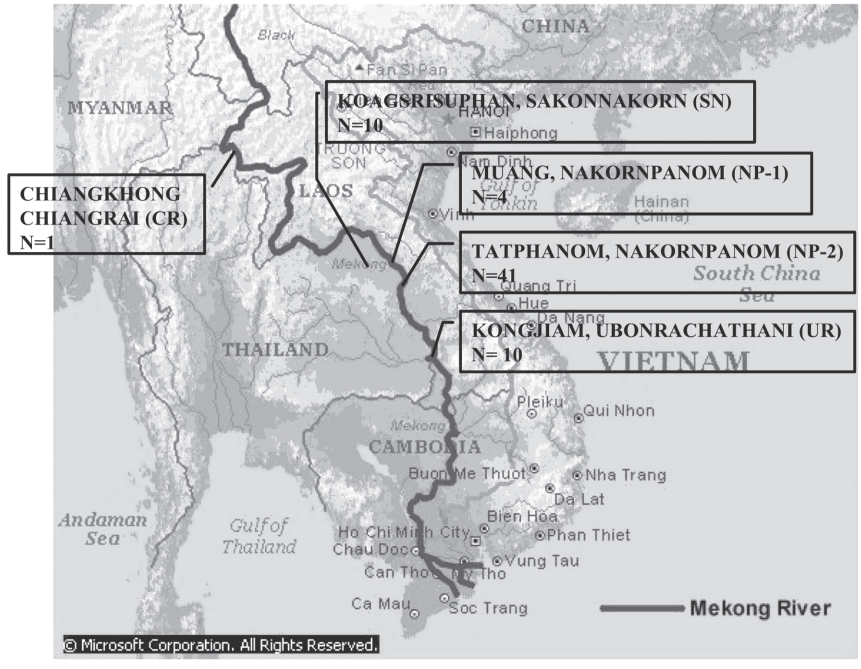

Figure 1. Map showing the five collection sites of *Pangasius sanitwongsei*

Scoring of Microsatellite Loci

All except one sample from Chiangrai (CR) were included in the scoring of microsatellite. DNA extraction was done using phenol/chloroform standard method (Taggart *et al*., 1992) with a slight modification. The individual DNA was re-suspended in TE buffer (10 mM Tris-HCl pH 7.2; 1 mM EDTA pH 8.0) and stored at 4°C. Primers for seven microsatellite loci developed from *P. gigas* DNA library, Pg1, Pg2, Pg3, Pg6, Pg9, Pg13 and Pg20 (detailed in Table 2) were used to amplify DNA samples in the polymerase chain reaction (PCR). Each 10 µl of PCR reaction contained 5-15 ng of DNA template, 0.5 µM of reverse and forward primers, reaction buffer (8.8 mM Tris-HCL pH 8.3 and 43.9 mM KCL), 1.2–1.5 mM $MgCl_2$, 100 µM of dNTPs mix and 0.2 unit of *Taq* Polymerase (Pharmacia Biotech.). The PCR cycles were comprised of 1 cycle at 94°C for 1 min, continued by 35 cycles of 94°C (30 sec), appropriate annealing temperature (30 sec) and 72°C (1 min), and then terminated by 1 cycle of 72°C for 1 min.

The PCR products were separated in polyacrylamide gel and visualized by silver staining. Sizes of alleles were determined according to an M13 sequence ladder.

Table 2. Nucleotide sequences of seven microsatellite primers used in this study

Loci	GenBank Accession no.	Primer sequence (5'-3')	Repeat units	TA (°C)	Product size range (bp)
*Pg-1**	AY364900	F:GGCCTGTCACAATGTGTATTGC R:GTCTGAGGTAGGCCTGTGAGGAG	$(CA)_{18}$	60	200-260
*Pg-2**	AY364901	F:TGTGTCTAATCTTGTCCGTGCTG R:TACTGTTGGACCAGACGTTCCTC	$(GA)_{14}N_{116}$ $(GT)_{11}$	60	249-311
*Pg-3**	AY364902	F:CCAGCCCACATTAGGTAGCATC R:ACTAAAAGGCCTGACCCTTAGC	$(GT)_{16}$	60	200-249
*Pg-6**	AY364903	F:CATCTCATGCCGTCATTACTGTG R:AAAGTCCTAAAGGCTGCTTCAAC	$(CA)_8CG$ $(CA)_6$	60	249-311
*Pg-9**	AY364904	F:TGTTGAAATGGTGATGTGTGACC R:CCTTGTCCCATAGGAGAAAGCAC	$(CT)_{11}$	55	200-249
*Pg-13**	AY364905	F:GTTTTCCATCCAGGTTGTTTTCC R: TAAGTCCATGTGGGTTTCCTCTG	$(CA)_{25}$	60	200-270
*Pg-20**	AY364909	F: TCCCACTACTCCCACCCACTTAC R: TATCTCTGATTTTGCACCACAGG	$(CT)_{13}$	60	180-220

Analysis of 16S rRNA

The 16S rRNA region of mt-DNA was amplified using the primer $16S_{F1}$ 5′ CGC CTG TTT AAC AAA AAC AT 3′ and $16S_{R1}$ 5′ CCG GTC TGA ACT CAG ATC ATG T 3′ (Palumbi *et al*., 1991). Each of a 30 µl reaction contained 1xPCR buffer, 2mM $MgCl_2$, 0.2 M dNTPs, 0.5 µM primer, 1 unit of *Taq*. polymerase and 50 ng DNA template. The reaction profile was as follows: a cycle of 94ºC, 3 min, 30 rounds of one minute each at 94, 52 and 72ºC; and a final round of 72ºC, 5 min. Then the PCR products were purified using ExoSAP-IT® (usb), sequenced using the Terminator Cycle Sequencing Ready Reaction Kit. The results were analyzed using ABI Prism® 377 DNA Sequencer (Applied Biosystems).

Data analyses

Microsatellite data: GENEPOP computer package (Raymond & Rousset, 1995) was used to calculate the number of alleles per locus, number of effective alleles per locus, observed and expected heterozygosities. Conformity to Hardy-Weinberg equilibrium was tested using Markov chain method (dememorisation: 1000, batches: 500 and iterations per batches: 1000) using the same package.

F-statistics (F_{IS}, F_{IT} and F_{ST}), test of genetic disequilibrium for each pair of locus and genetic differentiation between populations (dememorisation: 1000, batches: 100 and iterations per batches: 1000) were performed using the GENEPOP package. The overall F_{ST} was tested against 0 by bootstrapping. The locuswise F_{IS} for each population was also calculated to detect effects of inbreeding and Wahlund effects.

Cavalli-Sforza and Edwards (1967) chord distance was calculated. A neighbor-joining tree including bootstrap values (1000 replicates) was constructed using the PHYLIP computer package (Felsenstein, 1995).

The independent t-test comparison (Archie, 1985) was employed to test differentiation of observed heterozygosity among populations.

mt-DNA: Multiple alignments of the sequences were performed using the Clustal X program (Thompson *et al*., 1997). Then number of haplotypes and number of variation sites were enumerated. The genetic distance between haplotypes was calculated following Kimura (1980). A neighbor-joining tree (Saitou and Nei, 1987) was reconstructed and bootstrap values (1000 replicates) were calculated using the MEGA3 program (Kumar *et al*., 2004). Calculations of nucleotide diversity and F_{ST} among populations were facilitated by Arlequin ver 2.0 (Schneider *et al*., 2000).

RESULTS

Gene frequencies of four populations of *P. sanitwongsei* at seven microsatellite loci are shown in Table 3. Among seven microsatellite loci, *Pg6** was the most polymorphic with six alleles (2-6 alleles/population). *Pg3** and *Pg9** were equally polymorphic with four alleles across populations but the number of alleles/population was quite different (2-3 and 1-4 alleles/population for *Pg3** and *Pg9**, respectively).

Table 3. Allele frequencies at seven microsatellite loci of *Pangasius sanitwongsei* collected from four locations in the Mekong River basin. Population abbreviations as in table 1. N = sample size

Alleles	**NP-1**	**NP-2**	**UR**	**SN**
N	**4**	**41**	**10**	**10**
*Pg1**				
218	0.0000	0.0694	0.1667	0.0000
224	1.0000	0.9306	0.8333	1.0000
*Pg2**				
196	0.1250	0.0526	0.0625	0.0000
200	0.8750	0.9474	0.9375	1.0000
*Pg3**				
190	0.0000	0.0385	0.0556	0.0000
192	0.0000	0.0385	0.0000	0.7500
198	0.8750	0.9231	0.9444	0.2500
200	0.1250	0.0000	0.0000	0.0000
*Pg6**				
260	0.6250	0.4146	0.4000	0.9000
262	0.2500	0.0366	0.0500	0.1000
264	0.1250	0.4268	0.4500	0.0000
268	0.0000	0.0854	0.0500	0.0000
272	0.0000	0.0244	0.0500	0.0000
276	0.0000	0.0212	0.0000	0.0000
*Pg9**				
188	0.6250	0.3158	0.0000	0.0000
198	0.3750	0.6579	1.0000	1.0000
258	0.0000	0.0132	0.0000	0.0000
168	0.0000	0.0132	0.0000	0.0000
*Pg13**				
244	0.2500	0.5000	0.5556	0.1500
246	0.7500	0.5000	0.4444	0.8500
*Pg20**				
216	0.2500	0.1375	0.1667	0.0000
222	0.7500	0.8625	0.8333	1.0000

Hardy-Weinberg equilibrium (HWE) and genotypic disequilibrium

HWE was violated towards homozygote excess in one of four populations, NP-1 (F_{IS} was shown in Table 4). Genetic disequilibrium was significant between *Pg9** and *Pg13** (P = 0.049) for NP-2 and *Pg3** and *Pg13** (P = 0.034) for SN, however the genetic disequilibrium was not statistically significant after the Bonferroni correction was applied (α = 0.5/84).

Table 4. Genetic variability of four *Pangasius sanitwongsei* populations collected from four locations in the Mekong River basin. Population abbreviations as in table 1. Abbreviations: N = number of samples; A = average allele number across loci; A_e = average effective allele number across loci; H_O, H_e = observed and expected heterozygosity, respectively. NA= not available

Populations	**N**	**locus**	A	A_e	H_o	H_e	F_{IS}
NP-1	4	*Pg1**	1.000	1.000	0.000	0.000	NA
	4	*Pg2**	2.000	1.280	0.250	0.250	0.000
	4	*Pg3**	2.000	1.280	0.250	0.250	0.000
	4	*Pg6**	3.000	2.133	0.500	0.607	0.200
	4	*Pg9**	2.000	1.882	0.250	0.536	0.571
	4	*Pg13**	2.000	1.600	0.500	0.429	-0.200
	4	*Pg20**	2.000	1.600	0.000	0.429	1.000
Average			2.000 ^{ac} (0.577)	1.539 ^{a} (0.388)	0.250 ^{a} (0.204)	0.357 ^{a} (0.266)	0.333
NP-2	38	*Pg1**	2.000	1.148	0.139	0.131	-0.061
	38	*Pg2**	2.000	1.111	0.105	0.101	-0.042
	39	*Pg3**	3.000	1.170	0.154	0.147	-0.048
	41	*Pg6**	6.000	2.751	0.317	0.644	0.511
	38	*Pg9**	4.000	1.876	0.553	0.473	-0.170
	36	*Pg13**	2.000	2.000	0.555	0.507	-0.097
	40	*Pg20**	2.000	1.311	0.275	0.240	-0.147
Average			3.000 ^{a} (1.527)	1.624 ^{a} (0.615)	0.300 ^{a} (0.189)	0.321 ^{ac} (0.217)	0.066
UR	9	*Pg1**	2.000	1.385	0.333	0.294	-0.143
	8	*Pg2**	2.000	1.133	1.125	1.125	0.000
	9	*Pg3**	2.000	1.117	0.111	0.111	0.000
	10	*Pg6**	5.000	2.703	0.300	0.663	0.561
	8	*Pg9**	1.000	1.000	0.000	0.000	NA
	9	*Pg13**	2.000	1.976	0.444	0.523	0.158
	9	*Pg20**	2.000	1.385	0.333	0.294	-0.143
Average			2.286 ^{ac} (1.254)	1.528 ^{a} (0.609)	0.235 ^{a} (0.158)	0.287 ^{ac} (0.237)	0.189
SN	10	*Pg1**	1.000	1.000	0.000	0.000	NA
	10	*Pg2**	1.000	1.000	0.000	0.000	NA
	10	*Pg3**	2.000	1.600	0.300	0.395	0.250
	10	*Pg6**	2.000	1.219	0.200	0.189	-0.059
	9	*Pg9**	1.000	1.000	0.000	0.000	NA
	10	*Pg13**	2.000	1.342	0.300	0.268	-0.125
	10	*Pg20**	1.000	1.000	0.000	0.000	NA
Average			1.429 ^{bc} (0.534)	1.166 ^{a} (0.235)	0.114 ^{a} (0.146)	0.122 ^{bc} (0.163)	0.065
Across populations			3.1429 (1.5736)	1.6083 (0.5477)	0.2582 (0.1402)	0.3264 (0.1947)	

Note: Mean values in the same column with different superscripts are statistically significant (α = 0.05); A, A_e: Mann-Whitney test; H_O, H_e: *t*-test (Archie, 1985)

Genetic variation within population

Genetic variation of *P. sanitwongsei* is obviously low and did not vary much among the four populations included in this study. The number of alleles per locus ranged between 1.43 (±0.53) and 3.00 (±1.53). The effective number of alleles per locus was relatively low and ranged between 1.17 (±0.23) and 1.62 (±0.61). H_O was between 0.11 (±0.15) and 0.3 (±0.19) and H_e ranged from 0.12 (±0.16) to 0.36 (±0.27). The details on genetic diversity are shown in table 4. The Mann-Whitney test showed a significant difference only between the mean number of allele/locus of NP-2 and SN, while the *t*-test (Archie,1985) revealed no statistical differences between H_O or H_e of each population pair except for H_e of NP-1 and SN.

Population structuring and genetic diversity among populations

F_{st} was statistically significant (F_{st} = 0.190; $P < 0.01$) (Table 5) thus implying that substructure occurred among *P. sanitwongsei* populations. F_{is} (Table 5) was also significant implying that significant variations occurred within populations.

The locus-wise F_{is} did not show any locus specific nor population specific statistical significance (Table 6).

Table 5. F_{IS}, F_{ST} and F_{IT} values for each locus across all populations and average values across all loci based on Weir and Cockerham (1984)

Locus	F_{IT}	F_{ST}	F_{IS}
*Pg-1**	-0.051	0.031	-0.084
*Pg-2**	-0.050	-0.013	-0.037
*Pg-3**	0.577	0.554*	0.053
*Pg-6**	0.537	0.118*	0.474*
*Pg-9**	0.164	0.230*	-0.086
*Pg-13**	0.050	0.104*	-0.060
*Pg-20**	0.014	0.019	-0.005
All loci	0.276*	0.190*	0.106*

Note: An asterisk indicates statistical significance ($P < 0.01$).

Table 6. Locus wise F_{IS} values for each population

Locus	Populations			
	NP-1	NP-2	UR	SN
*Pg-1**	NA	-0.061	-0.143	NA
*Pg-2**	0	-0.042	0	NA
*Pg-3**	0	-0.048	0	0.2500
*Pg-6**	0.200	0.511*	0.561	-0.0590
*Pg-9**	0.571	-0.170	NA	NA
*Pg-13**	-0.200	-0.097	0.158	-0.1250
*Pg-20**	1.000*	-0.147	-0.143	NA

Note: NA= not available; * statistically significant ($P < 0.0071$)

Population differentiation, pairwise F_{ST} and genetic distance

The Fisher's exact test showed that SN was unique and significantly different from NP-1, NP-2 and UR ($P < 0.0083$, Bonferroni corrected) while the other population pairs were not different.

Pairwise, F_{ST} confirmed the difference between SN and other populations. In addition the difference between UR and NP-1 was also statistically significant under this analysis (Table 8).

Genetic distance ranged between 0.0952 and 0.4872.

Table 7. Pairwise F_{ST} between populations of *Pangasius saniwongsei* in Thailand

Population	NP-1	NP-2	UR
NP-1	***		
NP-2	0.0453^{ns}	***	
UR	0.1492*	0.0193ns	***
SN	0.3988*	0.3094*	0.3860*

Table 8. Genetic distance based on Cavalli-Sforza & Edwards (1967) between populations of *Pangasius saniwongsei* in Thailand

Population	NP-1	NP-2	UR
NP-1	***		
NP-2	0.0952	***	
UR	0.3120	0.1270	***
SN	0.4872	0.3014	0.2684

A phylogenetic dendrogram

The neighbor-joining tree showed that among the four populations of *P. sanitwongsei*, NP-2 was the most genetically unique and substantially diverged from NP-1 which came from a nearby location. SN was in between NP-2 and a group of UR and NP-1 with a 100% bootstrap value. UR and NP-1 were the most genetically similar population although the node was not supported by bootstrap value (52.3%) (Figure 2).

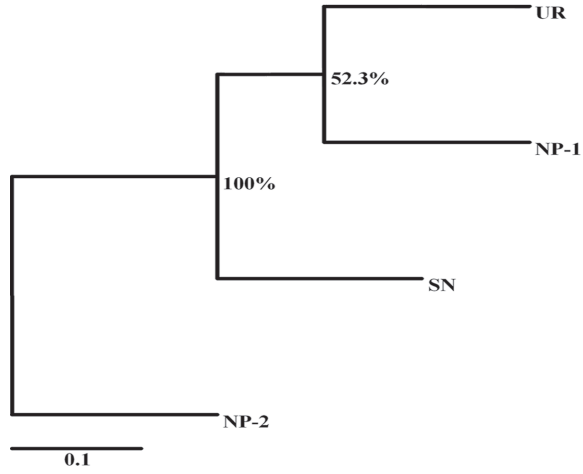

Figure 2. A phylogenetic dendrogram (neighbor-joining) based on genetic distance of four populations of *Pangasius sanitwongsei*. A number at each node is a bootstrap value

Diversity at 16SrRNA

Variation of 571 basepairs of 16SrRNA was not high which is common because this mt-DNA region is quite conserved (Hillis *et al*., 1996). However there was substantial variation whereby a total of six haplotypes were observed. PSh01 was common and found in all populations except for CR (n = 1). The PSh02 haplotype was observed in CR and SN while PSh03-06 existed in NP-2 only. Variation occurred including five transitional mutation (from purine to purine and pyrimidine to pyrimidine) and three transversional mutation (from purine to pyrimidine and *vice versa*) (Table 9).

Table 9. Summary of nucleotide variations found in the 16S rRNA fragment sequenced and haplotype frequencies in *Pangasius sanitwongsei*

Haplotype	Nucleotide sites								Haplotype frequency				
	89	157	383	390	414	480	512	536	CR (N=1)	NP-1 (N=4)	NP-2 (N=40)	SN (N=10)	UR (N=10)
PSh01	G	C	C	C	A	C	C	C	0	4	36	2	10
PSh02	.	.	.	T	.	.	.	.	1	0	0	8	0
PSh03	A	T	.	.	.	.	.	.	0	0	1	0	0
PSh04	.	.	G	.	.	A	.	.	0	0	1	0	0
PSh05	.	.	.	.	.	.	T	G	0	0	1	0	0
PSh06	.	.	.	.	G	.	.	.	0	0	1	0	0

Note: Only variable sites are shown. Dots indicate identity with the PSh01 haplotype sequence.
CR = Chiangrai, NP = Nakornpanom, SN = Sakonnakorn, UR= Ubonratchatani, N = sample size.

Among the six haplotypes observed in this study PSh03, 04 and 05 were the most distinct with two variable sites in each sequence. PSh03 contained only the transitional mutation PSh04 and 05 included one transversional mutation. The pairwise genetic distance between haplotypes showed highest distance among these three haplotypes while PSh01 and 02 were the most similar (Table 10).

Table 10. Number of haplotypes, variable sites, haplotype diversity and nucleotide diversity of five populations of *Pangasius sanitwongsei*

	PSh01	**PSh02**	**PSh03**	**PSh04**	**PSh05**	**PSh06**
PSh01	*****					
PSh02	0.002	*****				
PSh03	0.004	0.005	*****			
PSh04	0.004	0.005	0.007	*****		
PSh05	0.004	0.005	0.007	0.007	*****	
PSh06	0.002	0.004	0.005	0.005	0.005	*****

The overall haplotype diversity was 0.3452±0.0699 and nucleotide diversity was 0.0008±0.0008. It should be noted that sample sizes were varied (from 1 to 40) therefore any comparison between populations should be made with caution. Regarding each population the number of haplotypes varied from 1-5 haplotypes/population, the number of variable sites was between 0-7 sites, haplotype diversity ranged from 0.000 to 0.3556, and nucleotide diversity ranged between 0 to 0.0006 (Table 11).

Table 11. Pairwise genetic distance between haplotypes

Sample code	**No. of samples**	**No. of haplotype**	**No. of variable sites (V)**	**Haplotype diversity (h)**	**Nucleotide diversity (π)**
CR	1	1	-	1.0000±0.0000	0.0000±0.0000
NP-1	4	1	-	0.0000±0.0000	0.0000±0.0000
NP-2	40	5	7	0.1923±0.0831	0.0006±0.0007
	10	2	1	0.3556±0.1591	0.0006±0.0007
SK					
UR	10	1	-	0.0000±0.0000	0.0000±0.0000
Overall	65	6	8	0.3452±0.0699	0.0008±0.0008

Diversity of 16SrRNA within and between populations

Mean pairwise distance (Table 12) within populations was relatively high for NP-2 (0.005) followed by SN (0.002) and was nil for other populations. Despite the short physical distance between the collection sites, NP-1 and NP-2 showed substantial distance (0.0028).

Most of the pairwise F_{ST} was statistically significant except for the pair with CR because of very low sample size, and a pair between NP-1 and UR (Table 13). Figure 3 shows a phylogenetic dendrogram of the six haplotypes observed in this study.

Table 12. Mean pairwise distance within and between populations of *Pangasius sanitwongsei*

	CR	NP-1	NP-2	SN	UR
CR	0.0000				
	(0.0000-0.0000)				
NP-1	0.0020	0.0000			
	(0.0020-0.0020)	(0.0000-0.0000)			
NP-2	0.0042	0.0028	0.0050		
	(0.0020-0.0050)	(0.0000-0.0040)	(0.0020-0.0070)		
SN	0.0010	0.0010	0.0035	0.0020	
	(0.0000-0.0020)	(0.0000-0.0020)	(0.0000-0.0050)	(0.0020-0.0020)	
UR	0.0020	0.0000	0.0028	0.0010	0.0000
	(0.0020-0.0020)	(0.0000-0.0000)	(0.0000-0.0040)	(0.0000-0.0020)	(0.0000-0.0000)

Table 13. Pairwise F_{ST} between samples of *Pangasius sanitwongsei*. An asterisk represents statistical significance ($P < 0.01$)

	CR	NP-1	NP-2	SN	UR
CR	******				
NP-1	1.0000	******			
NP-2	0.8077	0.8316*	******		
SN	0.6444	0.7484*	0.7666*	******	
UR	1.0000	1.0000	0.8516*	0.8222*	******

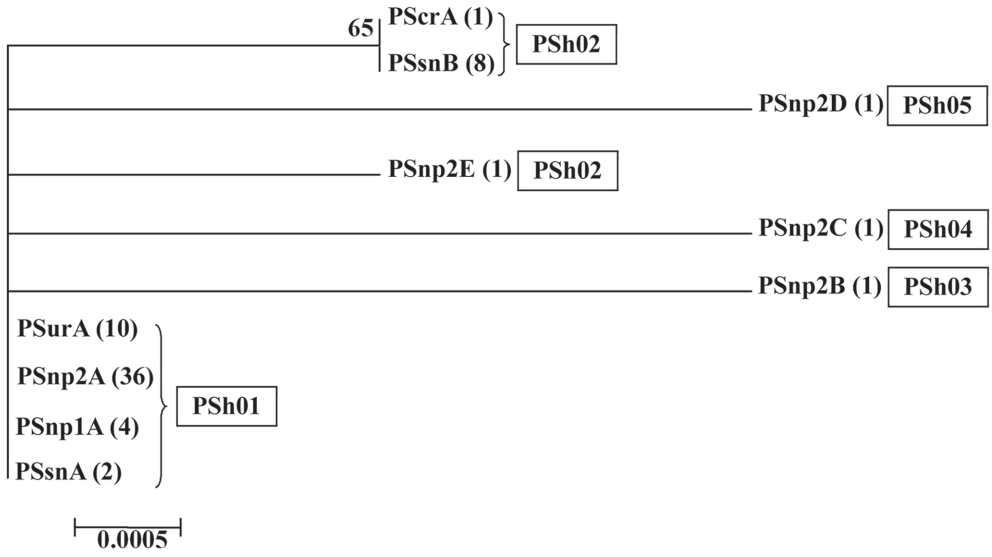

Figure 3. A phylogenetic dendrogram of six haplotypes observed in this study. Bootstrap values were 100%

DISCUSSION

Low genetic variation within populations

Genetic variation of *P. sanitwongsei* was obviously low hence indicating very low effective population size, which is common for endangered species (Frankham *et al*., 2002). For example, the northern hairy nosed wombat, *Lasiorhinus latifrons* had much lower genetic diversity than the more common and closely related southern hairy nosed wombat, *L. krefftii* ($A = 2.1$; $H_e = 0.32$ compared to $A = 5.9$; $H_e = 0.71$, respectively) (Taylor *et al*.,1994; Beheregaray *et al*.,2000). H_O of each population of *P. sanitwongsei* was significantly lower than that of critically endangered *P. gigas* collected in 2005 (N = 4, $H_O = 0.595 \pm 0.278$; $t_{\alpha.05, df\ 12} = 2.179$; Archie, 1985) based on the same microsatellite loci (Na-Nakorn *et al*., unpublished data). Similarly most populations had similar levels of H_O to the samples of seven *P. gigas* collected in 2004 ($H_O = 0.367 \pm 0.219$) (U. Na-Nakorn, unpublished data) ($P > 0.05$). The same trend was shown for H_e where no statistical difference was observed between these two species (H_e of *P. gigas* = 0.44 ± 0.24; 0.53 ± 0.21 for the 2004 and 2005 samples, respectively).

On the other hand, genetic variation of *P. sanitwonsei* was much lower than that of the common congeneric *Pangasius larnaudii*, based on four microsatellite loci ($H_O = 0.51 \pm 0.28$ - 0.82 ± 0.14; $H_e = 0.61 \pm 0.27$ – 0.79 ± 0.08 for *P. larnaudii*) (J. Fongisara, pers. comm.). A statistical test was not performed due to unequal number of loci used.

Owing to the close relationship of H_e and effective population size (Frankham *et al*., 2002), the results in this study suggested that *P. sanitwongsei* is under the same level of threat as being faced by *P. gigas*. Although it was likely that the census number of *P. sanitwongsei* might be greater than that of *P. gigas* as indicated by more available

samples of *P. sanitwongsei*, most of the cases' census number was much larger than effective number (N_e) of brooders (Frankham, 1996). Therefore, this study certainly suggests that *P. sanitwongsei* deserves reconsideration from IUCN for re-categorizing this species in the list (Hogan *et al.*, 2009).

Among the four populations, genetic variations of NP-1, NP-2 and UR were not different based on all parameters whereas SN showed a significantly lower number of alleles/locus and H_e than those of the other three populations. This reflects much smaller effective population size in SN. Due to the very small sample size of NP-1, it is unlikely to attribute the departure from HWE of NP-1 to demographic events (Hedrick, 1999).

Population structuring revealed by microsatellites

Substructure of *P. sanitwongsei* existed as revealed by $F_{ST} = 0.190$ ($P < 0.05$) which is relatively strong when compared to the congeneric migratory *P. larnaudii* ($F_{ST} = 0.099$) (J. Fongissara, pers. comm.) and *P. gigas* ($F_{ST} = 0.022$) (U. Na-Nakorn, unpublished). Substructure occurs when gene-flow is limited hence this data may indicate less extent of migratory behavior of *P. sanitwongsei* compared to these two species. However, the F_{ST} value obtained in this study was lower than the average degree of genetic subpopulation differentiation estimated for 49 species of freshwater fish ($G_{ST} = 0.222$) (Ward *et al.*, 1994). The strong substructure was supported by the high value of genetic distance which was obvious when compared to the common congeneric *P. larnaudii* (0.025 - 0.128) (J. Fongissara, pers. comm.).

The neighbor-joining tree separated NP-1 away from the genetically similar NP-2 (non-significant population differentiation test and pairwise F_{ST}) while the 16S rRNA showed that they shared a common haplotype. This might be a result of very small number of NP-1 thus biased the tree topology.

Diversity of 16S rRNA gene

The ratio of transitional mutation to transversions was 5:3 which is typical for variations within species, such as 5:1 and 3:1 in two *Mullus* species (Apostolidis *et al.*, 2001) and 7:5 in armored catfish (Shimabukuro-Dias *et al.*, 2004); 1:0, 2:0, and 3:1 in *Trachurus trachurus*, *T. mediterraneus* and *T. picturatus*, respectively (Karaiskou *et al.*, 2003). It is believed that the proportion of transversions increases with the increasing degree of genetic divergence (Apostolidis *et al.*, 2001).

Evolutionary Significant Unit

The Evolutionary Significant Unit (ESU) represents the genetically distinct populations within a species among which some of them may be on the way to speciation (Frankham *et al.*, 2002). Two evolutionary significant units were identified amongst the wild samples based on microsatellites, NP-1&2 and UR. The 16S rRNA genes gave slight different results for example NP-1 and NP-2 were significantly different which was possibly a bias due to small sample size of NP-1. Further the phylogenetic tree showed that the five haplotypes found in NP-2 were paraphyletic which might indicate more than one ESU in the samples (Frankham *et al.*, 2002). However such divergence included a small number of samples therefore we preliminarily drew a conclusion of two ESU.

The data from the two markers gave surprising results for SN. While it possessed very low genetic diversity at microsatellite loci it had relatively high haplotype and nucleotide diversity. We surmised that a major component of this stock could have been the offspring of wild fish bred in the fisheries station. The use of more than one female parent or the inclusion of a few wild samples (see description of the sample) could have been the explanation of the relatively high mt-DNA diversity.

Thus, we would suggest regarding SN as a hatchery stock which is valuable because it possessed a unique genetic pattern. Due to the sharing of the haplotype PSh02 with CR it was likely that this stock might have originated from a population from Chiangrai or its vicinity. This provides a clue for genetic distinctness of the CR populations, hence requiring further study.

CONCLUSION AND RECOMMENDATION

The present study validated the status of *P. sanitwonsei* in Thailand as an endangered species and thus requires prompt conservation efforts.

Population subdivision occurred even in the proximity of about 350 km along the Mekong River (between NP-2 and UR). Furthermore, there is a tendency that more genetically distinct populations will be discovered, for example in the Chiangrai population. Hence, further studies including extensive sampling are needed.

Based on the present study, it is recommended that NP-1 + NP-2 and UR should be independently managed.

REFERENCES

Apostolidis, AP, Mamuris, Z, Trantaphyllidis, C. 2001. Phylogenetic relationships among four species of Mullidae (Perciformes) inferred from DNA sequences of mitochondrial cytochrome b and 16S rRNA genes. **Biochem. Syst. Ecol.** 29, 901-909.

Archie, JW. 1985. Statistical analysis of heterozygosity data independent sample comparison. **Evolution** 39, 623-637.

Beheregaray, LB, Sunnucks, P, Alpers, DL, Banks, SC, Taylor, AC. 2000. A set of microsatellite loci for the hairy nosed wombats (*Lasiorhinus krefftii* and *L. latifrons*). **Conserv. Genet.** 1, 89-92.

Cavalli-Sforza, LL, Edwards, AWF. 1967. Phylogenetic analysis: models and estimation procedure. **Evolution** 21, 550-570.

Felsenstein, J. 1995. **PHYLIP (*Phylogeny Inference Package*) *Version 3.57c.*** Available Source: http://evolution.genetics.washington.edu/phylip.html. 1 January 2003.

Frankham, R. 1996. Relationship of genetic variation to population size in wildlife. **Conserv. Biol.** 10, 1500-1508.

Frankham, R, Ballou, JD, Briscoe, DA. 2002. ***Introduction to Conservation Genetics.*** Cambridge University Press, Cambridge.

Froese, R, Pauly, D. 2005. ***FishBase.*** *World Wide Web electronic publication.* www.fishbase.org, version (06/2005). Access date: 24th July, 2005.

Hedrick, PW. 1999. ***Genetics of Population, 2nd edn.*** Jones and Bartlett Publishers, Sudbury, Massachusetts.

Hillis, DM, Moritz, C, Mable, BK. 1996. ***Molecular Systematics 2nd edition.*** Sinauer Associates, Inc. Massachusetts, USA.

Hogan, Z, Na-Nakorn, U, Kong, H. 2009. Threatened fishes of the world: *Pangasius sanitwongsei* Smith 1931 (Siluriformes: Pangasiidae). **Environ. Biol. Fish.,** 1-2.

Jensen, JG.1996. ***1,000,000 tonnes of fish from the Mekong? Mekong Fish Catch and Culture,*** Mekong River Commission Secretariat, Phnom Penh, Cambodia.

Karaiskou, N, Apostolidis, AP, Trantaphyllidis, A, Kouvatsi, A, Trantaphyllidis, C. 2003. Genetic identification and phylogeny of three species of the genus *Trachurus* based on mitochondrial DNA analysis. **Mar. Biotechnol.** 5, 493-504.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. **J. Mol. Evol.** 16, 111-120.

Kumar, S, Tamura, K, Nei, M. 2004. *MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment.* **Brief. Bioinfo.** 5,150-163.

Miller, MP. 1997. ***Tools for population genetic analyses (TFPGA) 1.3: A window program for the analysis of allozyme and molecular population genetic data.*** Arizona: Northern Arizona University.

Mitchell, R, Braun, D. 2003. **Giant catfish critically endangered, group says.** National Geographic News. NATIONALGEOGRAPHIC.COM.

OEPP. 1997. ***Report of the Meeting on Present Status of Bioresources in Thailand.*** Bangkok: Office of Environmental Policy and Planning. Ministry of Science, Technology and Environment. (In Thai)

Palumbi, SR, Martin, AP, Romano, S, McMillan, WO, Stice, L, Grabowski, G. 1991. **The Simple Fool's Guide to PCR,** pp. 1-45. Department of Zoology, University of Hawaii, Honolulu.

Poulsen, AF. 2005. **Integration of fishers' Knowledge into research on a large tropical river basin, the Mekong River in Southeast Asia.** pp. 198-207, in Putting Fishers' Knowledge to Work: Conference Proceedings. Mekong River Commission, Lao PDR.

Pouyaud, L, Teugels, GG, Gustiano, R, Legendre, M. 2000. Contribution to the phylogeny of Pangasiid catfishes based on allozymes and mitochondrial DNA. **J. Fish Biol.** 56, 1509-1538.

Raymond, M, Rousset, F. 1995. *GENEPOP (ver. 1.2): a population genetics Software for exact test and ecumenicism.* **J. Hered.** 86, 248-249.

Roberts, TR, Vidthayanon, C. 1991. Systematic revision of the Asian catfish family Pangasiidae, with biological observations and descriptions of three new species. pp. 97-144, in **Proc. Acad. Nat. Sci. Phil., USA.**

Saitou, N, Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. **Mol. Biol. Evol.** 4, 406-425.

Schneider, S, Roessli, D, Excofier, L. 2000. ***Arlequin: A software for population genetics data analysis.*** Genetics and Biometry Lab, Department of Anthropology, University of Geneva.

Shimabukuro-Dias, CK, Oliveira, C, Reis, RE, Foresti, F. 2004. Molecular phylogeny of the armored catfish family Callichthyidae (Ostariophysi, Siluriformes). **Mol. Phylo. Evol.** 32, 152-163.

Smith, HM. 1945. ***The Fresh-water Fishes of Siam or Thailand.*** Government Printing Office, Washington, USA.

Taggart, JB, Hynes, RA, Prodöhl, PA, Ferguson, A. 1992. A simplified protocol for routine total DNA isolation from salmonid fishes. **J. Fish Biol.** 40, 963 -965.

Taylor, AC, Sherwin, WB, Wayne, RK. 1994. Genetic variation of microsatellite loci in bottlenecked species: the northern hairy nosed wombat *Lasiorhinus krefftii*. **Mol. Ecol.** 3, 729-731.

Thompson, JD, Gibson, TJ, Plewniak, F, Jeanmougin, F, Higgins, DG. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. **Nucleic Acids Res.** 24, 4876-4882.

Ward, RD, Woodwark, M, Skibinski, DOF. 1994. A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. **J. Fish Biol.** 44, 213-232.

Weir, BS, Cockerham, CC. 1984. Estimating *F*-statistics for the analysis of population structure. **Evolution** 38 (6), 1358-1370.