5. RESULTS

1. Species identification by using morphology and morphometrics

In this study, morphological features were used to identify two hundred and thirty-seven adult specimens from eight locations in Thailand. The morphological features of the rostrum, the distance of gastro-orbital crest between the gastro-orbital ridge and the hepatic spine, including the third maxilliped in adult males, were used to identify these white prawns.

In general, the wild animals often show a number of morphological distinct within the same species, including the white prawn. This study, the white prawns were divided according to our observations into five groups (Table 6).

1. M group (147 specimens from all locations) has a high and broad triangular rostrum (Figure 3A). The distance of gastro-orbital crest between the gastro-orbital ridge and the hepatic spine is absent or nearly absent (Figure 3A). The third maxilliped feature shows the distal segment about half as long as the second segment which bears a tuft of dense short hairs (Figure 8A). These features are the morphological identification of *P. merguiensis*.

2. I group (27 specimens from PKW, SKE, TDE, SRE and STW). The rostrum is a sigmoidal shape (Figure 3C). The distance of gastro-orbital crest between the gastro-orbital ridge and the hepatic spine is well present (Figure 3C). The third maxilliped feature shows the distal segment about as long as the second segment which bears a tuft of dense long hairs (Figure 8B). These morphological features identify *P. indicus*.

3. M/I group (9 specimens from RNW, STW, NKE and TRW). The rostrum is a sigmoidal shape (Figure 3C) as in *P. indicus*. The distance of the gastro-orbital crest between the hepatic spine and the gastro-orbital ridge is nearly absent (Figure 3A) similar to *P. merguiensis*. The third maxilliped feature shows the distal segment only about half as long as the second segment which bears a tuft of dense short hairs (Figure 8A) similar to *P. merguiensis*. The morphological features are therefore a mixing of *P. indicus* and *P. merguiensis*.

4. ND group (45 specimens from PKW, SKE, SRE, RNW, NKE and TRW). The rostrum is a high shape (Figure 3A) as in *P. merguiensis*. The distance of gastro-orbital crest between the gastro-orbital ridge and the hepatic spine is well present as in *P. silasi* and *P. indicus* (Figure 3B and 3C). Most specimens in this group were female, so we were unable to pinpoint the species (*P. merguiensis* or *P. indicus* or *P. silasi*) because the third maxilliped can not be used to identify in females.

5. S group (9 specimens from NKE), The third maxilliped character in males can be separated in *P. silasi* also and these were classified in the S group. In the rostrum is a moderate curve (Figure 3B). The distance of gastro-orbital crest between the hepatic spine and the gastro-orbital ridge is well present (Figure 3B). The third maxilliped shows the distal segment about as long as the second segment which bears only a rudimentary tuft of hairs at a tip (Figure 8D). The morphological features are similar to *P. silasi*. However, the third maxilliped could not be used to identify species in females. Therefore, it is possible that there are mixed species in the ND group. To identify species more clearly, the morphometric characters were used.

Table 6 The five groups of 237 samples from all locations identified by morphological features

The ratio of L_1:$L_{1,2}$ of two hundred and thirty-seven specimens used to divide them into four groups (Table 7).

1. M group (199 specimens from location). The ratio of L_1: $L_{1,2}$ ranged from 0.27-1.63. This range value is broader than in Pendrey et al. (1999) reported (0.5-1.644). This group seems to classify as *P. merguiensis*.
2. I group (15 specimens derived from PKW, SKE, TDE, RNW, STW, NKE and TRW). The ratio of L_1: $L_{1,2}$ ranged from 2.0 to 2.88. This range corresponds to Pendrey et al. (1999), who reported 2.0-3.98 (L_1: $L_{1,2}$), which classifies as *P. indicus*.
3. M/I group (14 specimens from TDE, SRE, RNW, STW, NKE and TRW). The ratio of L_1: $L_{1,2}$ ranged from 1.7 to 1.88, not in the range of *P. merguiensis* or *P. indicus* as classified in Pendrey et al. (1999). This group could not be classified as to species.
4. In this study, in *P. silasi* were The L_1: $L_{1,2}$ value was between the values of M and I group (1.63-2.0) including the overlapping values in I (2.0-2.4). For the morphometric measurements of *P. silasi*, three specimens were in M/I group (NKE2, NKE17 and NKE19) and six specimens in I group (NKE5, NKE8, NKE12, NKE22, NKE27 and NKE31). Therefore, the morphometric characters could not be used to separate *P. silasi* from *P. indicus*.

Table 7. The statistical summary of morphometric measurements (L_1:$L_{1,2}$) of 237 specimens from eight locations in Thailand.

Group	L_1:$L_{1,2}$	Sample location (No.)
Morphometric as M (199) ***P. merguiensis***		PKW (15), SKE (31), TDE (23), SRE (26), RNW(32), STW (27), NKE (16), TRW (29)
Average ± SD	1.02 ± 0.33	
Maximum	1.63	
Minimum	0.27	
Morphometric as M/I (14) ***P. merguiensis*/*P. indicus***		TDE (5), SRE (1), RNW (1), STW (2), NKE (4), TRW (1)
Average ± SD	1.76 ± 0.06	
Maximum	1.88	
Minimum	1.7	
Morphometric as I (15) ***P. indicus***		PKW(1), SKE (3), TDE (2), RNW (1), STW (2), NKE (2), TRW (4)
Average ± SD	2.20 ± 0.26	
Maximum	2.88	
Minimum	2.0	
Morphology as *P. silasi* (9)		NKE (9)
Morphometric measurement		
Average ± SD	1.95 ± 0.26	
Maximum	2.4	
Minimum	1.63	

Finally, after using the morphological and morphometric characters to identify the specimens, the data were compared and categorized into three groups (Table 8).

Group I:

The morphological and the morphometric character provided a clearly correspondent identification as *P. merguiensis* (M group). There were 127 specimens in this group from all locations.

Group II:

The morphological and the morphometric characters provided a contrasting identification. The morphological characters led to identification as *P. merguiensis* (M group) while the morphometric characters led to *P. indicus* (I group). In the other hand, the morphological characters led to *P. indicus* (I group) while the morphometric characters led to *P. merguiensis* (M group). There were 41 specimens in this group from all locations.

Group III:

The morphological and morphometric characters could not assistant clearly to identify the species origin. The morphological characters could be classified in either the M/I group or the ND group, and the morphometric characters classified as the M/I group. There were 69 specimens in this group from all locations.

Table 8. Two hundred and thirty seven samples divided into 3 groups based on the comparison of morphological and morphometric data.

Group	Species according to morphology	Species according to morphometric	Sample location (No.)
I (127)	M	M	PKW(2), SKE (6), TDE (16), SRE (19), RNW (25), STW(25), NKE (10), TRW (24)
II (41)	M	I	PKW (1), SKE (3), TDE(2), RNW(1), STW(1), TRW(2)
	I	M	PKW(2), SKE (15), TDE(8), SRE (1), STW(1)
	S	I	NKE (4)
III (69)	ND	M	PKW(11), SKE (10), SRE (6), RNW(4), NKE (5), TRW (3)
	M	M/I	TDE(4), SRE (1), RNW (1), STW (2), NKE (1), TRW (1)
	M/I	M	RNW (3), STW (1), NKE (1), TRW (2)
	M/I	I	STW(1)
	ND	I	NKE (2), TRW (2)
	S	I	NKE (2)
	M/I	M/I	NKE (1)
	ND	M/I	NKE (2)
	S	M/I	NKE (3)

Morphology: M = *P. merguiensis*, I = *P. indicus*, S = *P. silasi*, M/I = mixed features of *P. merguiensis* and *P. indicus*, ND = species not determined

Morphometric: M = *P. merguiensis*, I = *P. indicus*, M/I = the L_1: $L_{1,2}$ ratio was between *P. merguiensis* and *P. indicus*.

PKW = Phuket, RNW = Ranong, TRW = Trang, STW = Satun, TDE = Trad, SRE = Surat Thani, NKE = Nakhon Si Thummarat, SKE = Songkhla.

2. Species identification by using allozyme pattern

Fifty-nine specimens (Table 9) from three groups (I, II and III) were assayed by MDH following Pendrey et al. (1999). Specimens from group I were morphologically and morphometrically identified as *P. merguiensis*. Group II samples were morphologically and morphometrically inconsistent. Group III samples were not morphologically or morphometrically determined. Male *P. silasi,* which has been successfully identified by third maxilliped, were also chosen for study of MDH patterns as an outgroup .

OF fifty-nine specimens from groups I, II and III, fifty-eight specimens and one specimen of *P. silasi* expressed a monomorphic band (two-banded pattern) for MDH (Figure 14), but one specimen (SKEB5) and five specimens of *P. silasi* expressed a polymorphic band (three-banded). This indicates that there is a low level of different patterns for MDH system to identify the different morphological and morphometric data. And that there are overlapping MDH patterns between *P. silasi* and *P. merguiensis*. Therefore, an MDH system cannot be used to separate *P. merguiensis* from *P. silasi*. Our results are similar to Sodsuk and Sodsuk (1999), who found that the MDH did not provide the polymorphic band in the populations of *P. merguiensis* in three locations in Thailand. However, Pendrey et al. (1999) could clearly separate *P. indicus* from *P. merguiensis* by MDH (Figure 15).

Because of the limited differences in this study, further investigation is needed. Because other isozyme studies (Limcharearn, 1997; Sheatun, 1997) have not distinguished the white prawn and low genetic diversity (Mulley and Latter, 1980), we targeted them into further DNA-based methods. Other studies (DeSalle et al.,

1987; Karl and Avise, 1992) have been limited to comparing proteins or allozymes (homogeneous allozyme pattern), but DNA data usually provide greater resolution.

Table 9. Sixty-five specimens from eight locations around Thailand identified to species according to morphology and morphometric measurement were chosen for isozyme study. ☐ = The morphological and morphometric specimens (59) in Group I, II and III. ▒ = The morphologically identified *P. silasi* specimens (6).

Sample No.	Sex	Species according to morphometrics	Species according to morphology	Group
PKW10	M	M	M	I
PKW15	F	M	M	I
SKE11	F	M	M	I
SKE27	M	M	M	I
TDE1	F	M	M	I
TDE12	M	M	M	I
TDE14	M	M	M	I
TDE22	M	M	M	I
TDE23	F	M	M	I
TDE47	M	M	M	I
SRE8	M	M	M	I
SRE14	M	M	M	I
RNW19	M	M	M	I
RNW27	F	M	M	I
RNW50	M	M	M	I
STW4	M	M	M	I
STW9	F	M	M	I
NKE1	M	M	M	I
NKE9	M	M	M	I
NKE13	M	M	M	I
NKE21	M	M	M	I
NKE24	M	M	M	I
TRW7	M	M	M	I
PKW13	M	M	I	II
PKW1	M	I	M	II
PKW7	M	M	I	II
SKE26	M	M	I	II
SKE36	F	I	M	II

Table 9 (continued)

Sample No.	Sex	Species according to morphometric	Species according to morphology	Group
SKE43	M	I	M	II
SKE92	M	M	I	II
RNW9	F	I	M	II
STW31	M	M	I	II
NKE5	M	I	S	II
NKE8	M	I	S	II
TRW11	F	I	M	II
TRW15	M	I	M	II
SKEB1	F	M	ND	III
SKEB5	F	M	ND	III
PKW16	F	M	ND	III
SKE96	F	M	ND	III
TDE2	M	M/I	M	III
TDE21	M	M/I	M	III
SRE9	M	M/I	M	III
SRE12	F	M	ND	III
SRE30	F	M	ND	III
RNW55	M	M	M/I	III
RNW60	F	M	M/I	III
RNW66	M	M/I	M	III
STW17	M	M/I	M	III
STW23	M	M	M/I	III
NKE2	M	M/I	S	III
NKE3	M	M/I	M/I	III
NKE14	F	I	ND	III
NKE17	M	M/I	S	III
NKE18	F	I	ND	III
NKE19	M	M/I	S	III
NKE23	M	M/I	M	III
TRW21	M	M	M/I	III
TRW22	M	M/I	M	III
TRW25	F	M	M/I	III
SREE1	F	-	M	-
SREE2	M	-	M	-
STWA5	F	-	I	-
STWA6	M	-	I	-
STWA7	F	-	I	-

Morphology: M = *P. merguiensis*, I = *P. indicus*, S = *P. silasi*, M/I = feature is *P. merguiensis* and *P. indicus*, ND = species not determined.

Morphometric: M = *P. merguiensis*, I = *P. indicus*, M/I = the L_1: $L_{1,2}$ ratio between *P. merguiensis* and *P. indicus*.

Group I = The morphological and morphometric character provide correspondent identification; II = The morphological and the morphometric characters provided a contrasting identification; III = The morphological and morphometric characters could not assistant clearly to identify the species.

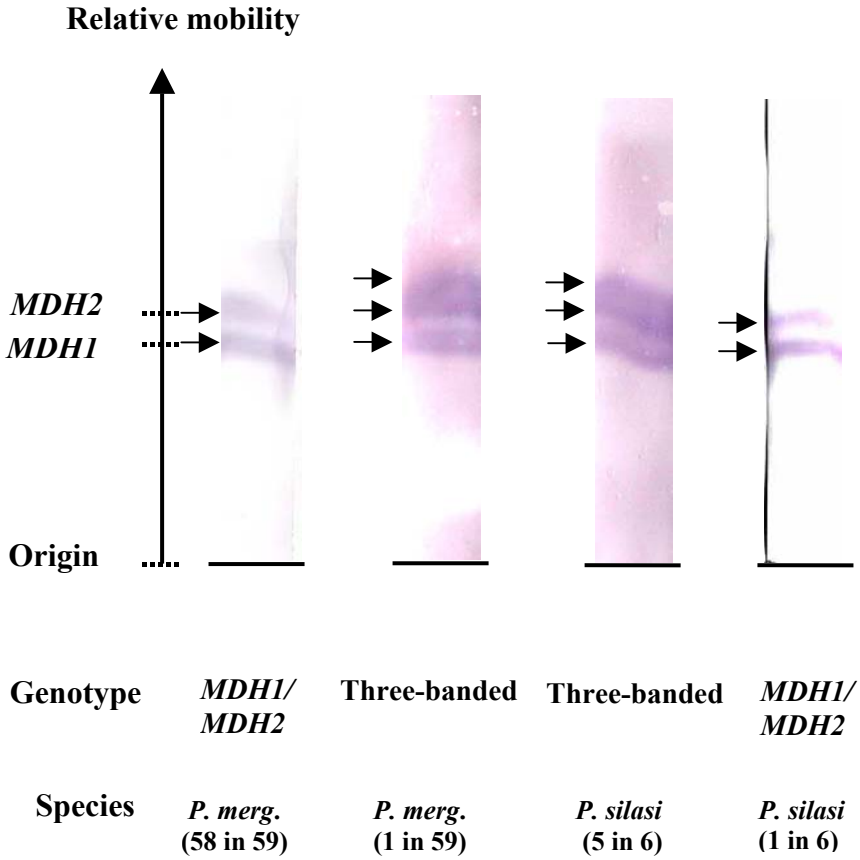

Figure 14. The MDH (Malate dehydrogenase) allozyme patterns in 65 samples of *Penaeus merguiensis (P. merg)* and *P. silasi* from both coasts of Thailand.

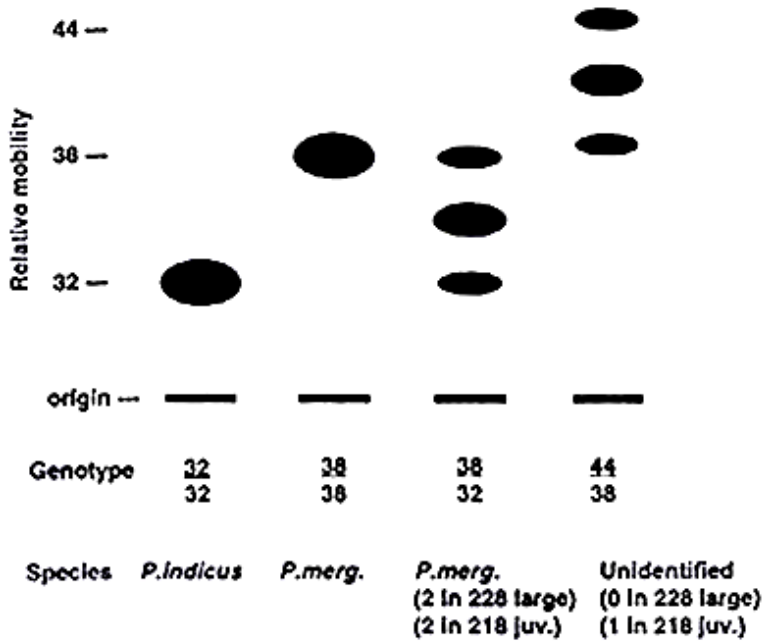

Figure 15. The MDH (Malate dehydrogenase) allozyme patterns of *P. merguiensis* (*P. merg.*) and *P. indicus* (Pendrey et al., 1999).

3. Species identification by using DNA markers

In this study, fourteen universal primer pairs (Table 10) were used to amplify DNA from the specimen mentioned earlier. Seven primer pairs (mtD-4, mtD-9; mtD-8, mtD-12; mtD-26, mtD-30; mtD-26, mtD-34; mtD-29, mtD-34; mtD-32, mtD-34; mtD-35, mtD-36) could be amplified and other primer pairs could not. The seven fragments obtained from amplification were compared with the GenBank database and were part of five genes in mitochondrial DNA (COI, Cytochrome oxidase subunit I; CB, Cytochrome b; ND1, NADH dehydrogenase subunit 1; 16S, the large subunit ribosomal RNA; and 12S, the small subunit ribosomal RNA). The first and second fragments were about 800 and 900 bp in length, and covered the COI gene. The third fragment was about 1600 bp in length and covered partial parts of the CB and ND1 genes. The fourth fragment was about 2400 bp in length and covered partial parts of the CB and 16S genes and we expected that it would cover the ND1 also (Figure 16). The fifth fragment was about 800 bp in length and covered partial parts of the ND1 and 16S genes. The sixth fragment was about 500 bp in length and covered a partial part of the 16S gene. The seventh fragment was about 400 bp in length and covered a partial part of the 12S gene.

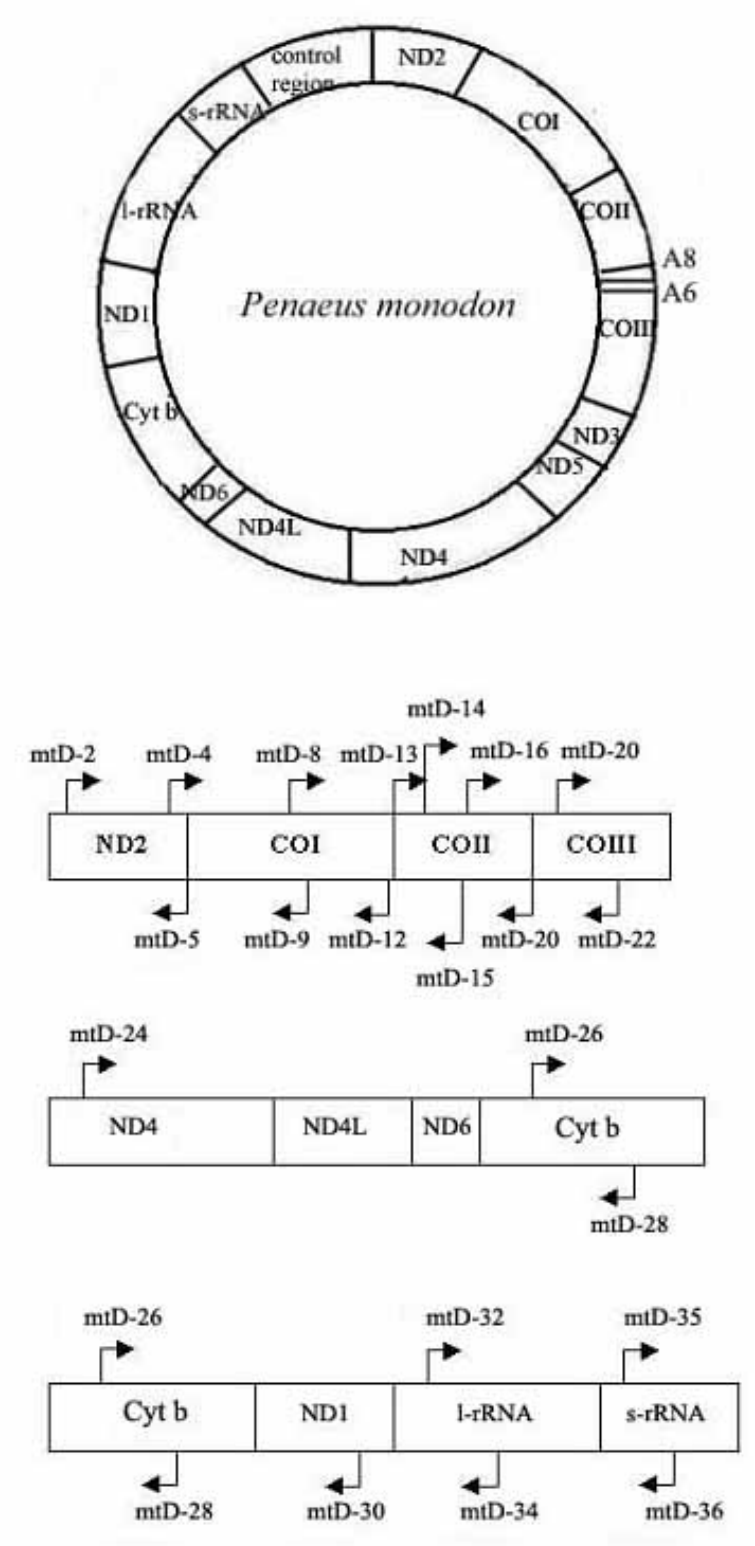

Figure 16. The mitochondrial genome of *P. monodon* and the direction of universal mitochondrial primer pairs (Simon et al., 1994).

Table 10. The amplification of the invertebrate universal primers (Simon et al., 1994).

Name	Region in genome	Amplification fragment
mtD-2, mtD-5	ND2	-
mtD-4, mtD-9	COI	~800bp
mtD-8, mtD-12	COI	~900bp
mtD-13, mtD-15	COII	-
mtD-13, mtD-20	COII	-
mtD-14, mtD-20	COII	-
mtD-16, mtD-20	COII	-
mtD-21, mtD-22	COIII	-
mtD-24, mtD-28	Cyt b, ND4	-
mtD-26, mtD-30	Cyt b, ND1	~1.6 kb
mtD-26, mtD-34	Cyt b, ND1, l6S	~2.4 kb
mtD-29, mtD-34	ND1, 16S	~800bp
mtD-32, mtD-34	16S	~500 bp
mtD-35, mtD-36	12S	~400 bp

COI = Cytochrome oxidase subunit I, COII = Cytochrome oxidase subunit II, COIII = Cytochrome oxidase subunit III, Cyt b = Cytochrome b, ND1 = NADH dehydrogenase subunit 1, ND2 = = NADH dehydrogenase subunit 2, ND4 = = NADH dehydrogenase subunit 4, 12S = The small subunit ribosomal RNA and 16S = The large subunit ribosomal RNA. - = no amplification.

To confirm the species status, this study attempted to find a *P. indicus* sample from another area (South Africa) as a reference. The sequence of *P. indicus* from this other area and the sequences of STWA6 and SREE1 were compared and then a phylogenetic tree constructed. In theoretical, STWA6 should be clustered in the same group of *P. indicus* and SREE1 should separate from *P. indicus*. Then we can be confident in species status.

Therefore, five specimens with firm in identification (Table 11) were used to represent each species for 12S and 16S amplification and the nucleotide sequencing. First, STWA6, a sample strongly identified as *P. indicus*; second, SREE1, a sample strongly identified as *P. merguiensis*; third, NKE8, a sample strongly identified as *P. silasi*; fourth, TA5, the DNA sample of *P. indicus* received from South Africa, and fifth, D1, the sample of *P. monodon* from Songkhla.

Table 11. The five samples chosen for partial sequence analysis of 12S and 16S gene.

Sample name	Character/Identification
STWA6	Strongly morphologically identified as *P. indicus*
SREE1	Strongly morphologically identified as *P. merguiensis*
NKE8	Strongly morphologically identified as *P. silasi*
TA5	DNA sample of *P. indicus* received from South Africa
D1	Strongly morphologically identified as *P. monodon*

3.1 12S rRNA

To clarify the species of STWA6, we cloned the seventh fragment covered partial part of the 12S gene and sequenced it. The 12S sequence data was located on the third domain of 12S gene (conserved primers SR-J-14233 (12Sbi) and SR-N-14588 (12Sai)).

The percent homology between SREE1 and others was 96.5 (STWA6), 98 (NKE8), 93 (TA5) and 89 (D1). In BLAST search results, all three species sequences (NKE8, TA5, and SREE1) were significantly matched to 12S rRNA sequence of *P. monodon* in 90% (Accession number: AF217843), *P. notialis*, 88% (Accession number: X84350) and *P. vannamei,* 87% (Acession number: S65259).

The length variation was also found in 12S rRNA domain III of insects (Page et al., 2002) and Lycosoidae spiders (Fang et al., 2000). This study, the fragment lengths were the same sizes (407 bp) in STWA6, SREE1, NKE8 and TA5 while 409 bp was found in D1. This 12S rRNA size corresponded to *P. monodon* (Accession number: AF217843) in GenBank database.

The %AT contents of STWA6, SREE1, NKE8, TA5 and D1 were 69.78, 69.78, 69.78, 68.06 and 70.41, respectively (Table 12). Our A-T content as found in 12S rRNA was consistent with descriptions of other invertebrates, i.e. 70% in Decapod crustacean (Ballard et al., 1992) and 70% in centipede (Ballard et al., 1992).

Table 12. The base composition percentages of 12S rRNA in STWA6, SREE1 (*P. merguiiensis*), NKE8 (*P. silasi*), TA5 (*P. indicus*) and D1 (*P. monodon*).

	%A	%T	%C	%G	%A+T
STWA6	34.89	34.89	17.20	13.02	69.78
SREE1	34.89	34.89	17.20	13.02	69.78
NKE8	34.89	34.89	17.20	13.02	69.78
TA5	34.15	33.91	18.43	13.51	68.06
D1	35.45	34.96	16.87	12.71	70.41

In the 12S sequence analyses, there were 56 substitution sites and more transitions than transversion (39:12). Transitions from C-T were commonly found more than A-G. The A-T transversions predominated over other transversions such as insect mitochondrial DNA (Simon et al., 1994). Figure 17 illustrates the insertion or deletion of nucleotide sequences and 11 informative parsimonious sites. Many positions showed a specific nucleotide sequence in each species. Table 13 shows four regions (84-89, 98-102, 144-148 and 188-191) of the variable positions in *P. merguiensis*, *P. silasi, P. indicus* and *P. monodon*.

Table 14, at the second column from the left hand, shows the genetic distance between SREE1 (*P. merguiensis*) and other species, NKE8 (*P. silasi*), TA5 (*P. indicus*) and D1 (*P. monodon*) is 0.0200 (2%), 0.0592 (5.92%) and 0.1201 (12.01%), respectively. *P. merguiensis, P. indicus and P. silasi* are closely related, and *P. monodon* is clearly differentiated. This result corresponds to morphological identification. The percent distance between STWA6 and TA5, at the fourth column and the fifth row in Table 14, was 7.78%, which is too distant to classify within species. In contrast, the percent distance between STWA6 and SREE1, at the second

column and the third row in Table 14, was 3.52 %, close together. The phylogenetic tree of the 12S-sequence (Figure 18) shows STWA6 close to SREE1 (*P. merguiensis*) and NKE8 (*P. silasi*), while they are both separate from TA5 (*P. indicus*). It seems that species identification of STWA6 by using the molecular data (12S) and morphological data are incongruent. STWA6 required more evidence to confirm the species status.

Table 13. The nucleotides occurring at variable position in 12S rRNA mtDNA gene of STWA6, SREE1 (*P. merguiensis*), NKE8 (*P. silasi*), TA5 (*P. indicus*) and D1 (*P. monodon*).

Sample name	First region (84-90)	Second region (98-102)	Third region (144-148)	Fourth region (188-191)
STWA6	TATAATA	ATTTA	GTATT	CTAG
SREE1	TATAACG	ATCTG	CTACT	TTAG
NKE8	TATAACA	ATTTG	CTACT	ATAG
TA5	ATCAACA	ATCCG	TTACT	TTAG
D1	AATAATA	CTTTA	TTATC	TCAA

Table 14. The distance matrix of 12S sequence of STWA6, SREE1 (*P. merguiensis*), NKE8 (*P. silasi*), TA5 (*P. indicus*) and D1(*P. monodon*).

	SREE1	NKE8	STWA6*	TA5	D1
SREE1	0.0000				
NKE8	0.0200	0.0000			
STWA6	0.0352	0.0200	0.0000		
TA5	0.0592	0.0645	0.0778	0.0000	
D1	0.1201	0.1144	0.1114	0.1304	0.000

* COI sequence was classified as *P. merguiensis*.

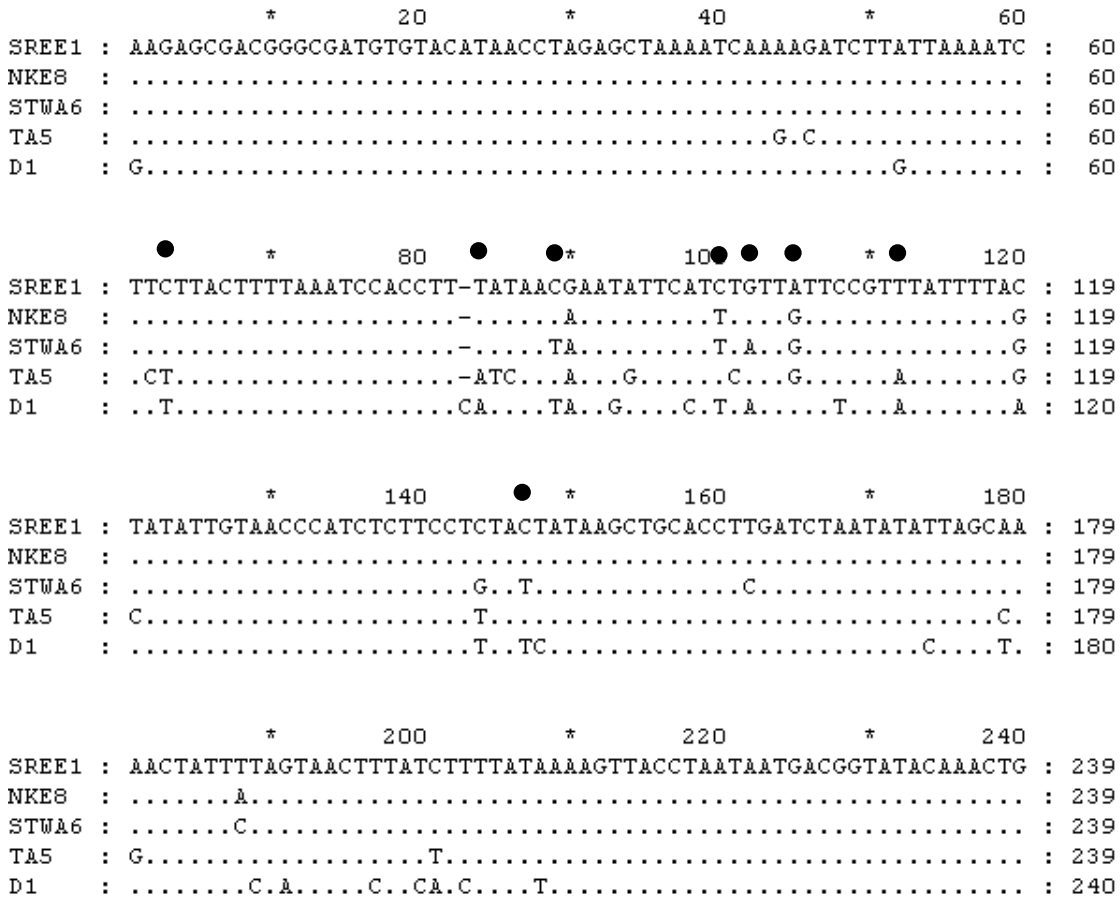

Figure 17. The nucleotide sequence alignments of the 12S rRNA genes in SREE1 (*P. merguiensis*), STWA6, NKE8 (*P. silasi*), TA5 (*P. indicus*) and D1 (*P. monodon*). ● Indicates informative sites. – Indicates the deletion position.

Figure 17. (continued)

```
            ●    *      ●260         *       280         *       300
SREE1 : TA-TTTCTACAAAATAGAGTAAGATTCTTCGTGGACTATCGATTACAGGACAGGTTCCTC : 298
NKE8  : ..-.................................................... : 298
STWA6 : ..-.................................................... : 298
TA5   : ..-.A........G..A...................................... : 298
D1    : ..A.AAT.........A........AT.......TT.........T.............. : 300

                  *       320         *       340         *       360
SREE1 : TAAATAGACTAAGTTACCGCCAAATCCTTTGAGTTTCAAGATAATAACTGTTTAGTACCC : 358
NKE8  : ..................................................... : 358
STWA6 : ..................................................... : 358
TA5   : ..........................................C.......... : 358
D1    : ............A.............................................. : 360

                  *    ●  380         *       400
SREE1 : AGGTAATAACAATTCAAATAAAGTATAATAGGGTATCTAATCCTAGTTT : 407
NKE8  : ..............T.................................. : 407
STWA6 : ..................................................... : 407
TA5   : ..............T.................................. : 407
D1    : .......T.T....T.........G...C...............G...C : 409
```

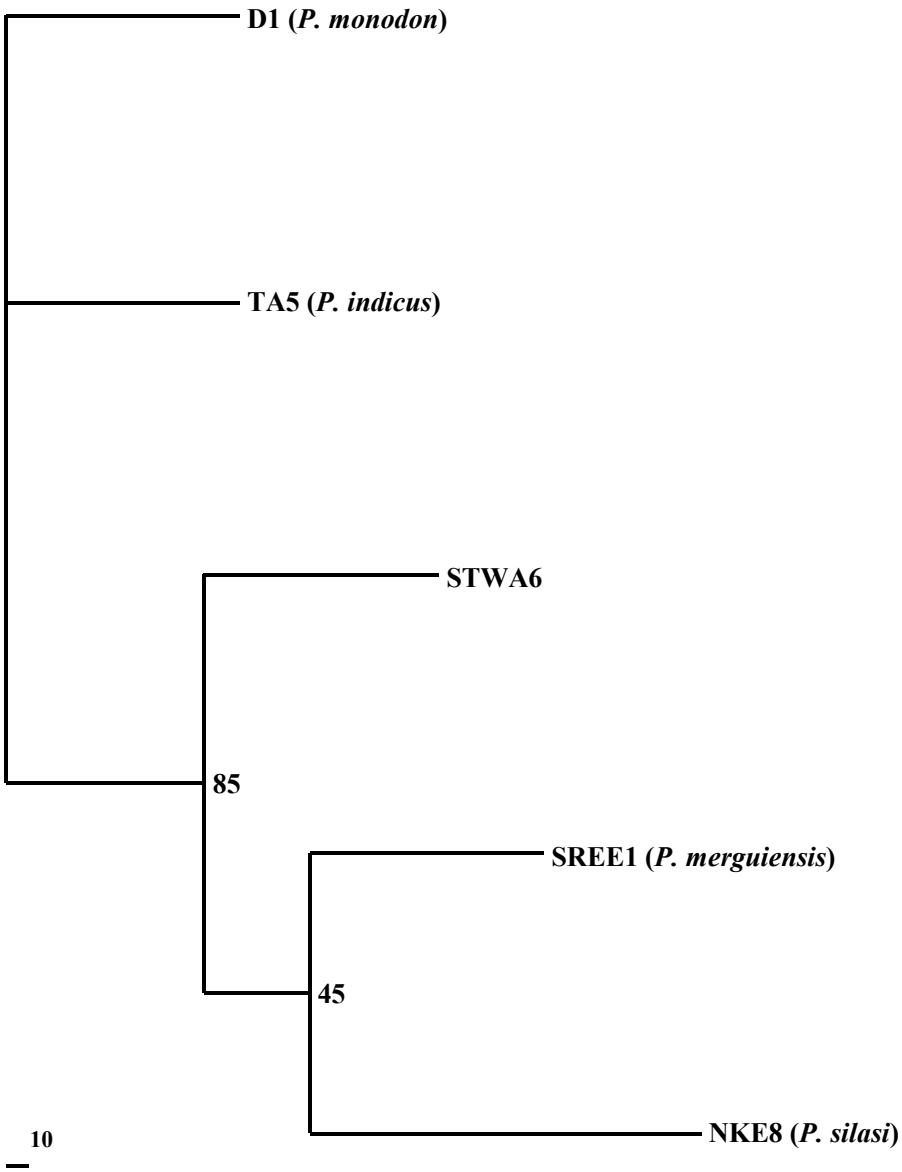

Figure 18. The neighbor-joining tree of 12S nucleotide sequences of sample STWA6, SREE1, NKE8, TA5 and D1. — Indicates the scaled branches phylogenetic tree: lengths of branches are proportional to the number of molecular changes.

3.2 16S rRNA

To specify the species status of STWA6, we cloned and sequenced the sixth fragment, which covered partial part of the 16S gene. The 16S sequence that we amplified is located on the first domain of the 16S gene (conserved primers LR-J-12887, LR-N-13398).

In BLAST search results, all three species sequences (NKE8, TA5, and SREE1) are identity to 16S rRNA sequence of *P. chinensis* (Accession number: AF217843) in 98 %, 96% and 96%, respectively.

The 16S rRNA sequences from STWA6, SREE1 (*P. merguiensis*), NKE8 (*P. silasi*), TA5 (*P. indicus*) and D1 (*P. monodon*) were 561, 561, 561, 560 and 562 bp, respectively. The %AT contents of STWA6, SREE1, NKE8, TA5 and D1 were 66.49, 66.13, 66.25, 65.96 and 67.79, respectively (Table 16). A high A-T content of the fragment under study is correlated with that of other invertebrates, for example, 76% in *Drosophila* (Clary and Wolstenholme, 1985), 67.7 % in *P. notialis* and 68.1% in *P. schmitti* (Machado et al., 1993). An increase of A-T content of mtDNA seems to be a general tendency in the evolution of arthropods (Zehethofer and Sturmbauer, 1998).

The percent homology between SREE1 and others was 97 (STWA6), 97 (NKE8), 98 (TA5) and 92 (D1), indicating that STWA6, SREE1, NKE8 and TA5 were close species similar to our 12S data above.

Table 15. The base composition percentages of 16S rRNA in STWA6, SREE1 (*P. merguiensis*), NKE8 (*P. silasi*), TA5 (*P. indicus*) and D1 (*P. monodon*).

	%A	%T	%C	%G	%A+T
STWA6	33.51	32.98	19.79	13.73	66.49
SREE1	32.80	33.33	19.79	14.08	66.13
NKE8	33.16	32.80	19.96	14.08	65.96
TA5	33.21	33.04	20.00	13.75	66.25
D1	34.52	33.27	18.86	13.35	67.79

There were 50 variable sites in the sequence alignment, including four insertion and deletion positions (10, 187, 217 and 499) (Figure 19). There was transition more than transversion (29:17). Transitions from C-T were commonly found more than A-G, similar to our 12S data. The A-T transversions predominated over other transversions, as in insect mitochondrial DNA (Simon et al., 1994) and 12S of our results above. There was one informative parsimonious site (287), as shown in Figure 19. As a result it should be constructed the phylogenetic tree by the distance method instead of the parsimony method. There were many positions that showed specific nucleotide sequences in each species. Twenty-nine positions of D1 (*P. monodon*) exhibited specific nucleotide sequences and could be differentiated from the other species. The positions 267, 268, 387, 388 and 554 in NKE8 (*P. silasi*) illustrate the differentiation from other species. In TA5 (*P. indicus*), the positions 44, 196, 267, 268, 381 and 535 provided distinct nucleotide from other species. At the position 267-268, the same base (TT) was found in SREE1 and STWA6, while different nucleotide sequences were found in NKE8 (TC), TA5 (TA) and D1 (CA) (Table 16).

The genetic distance (Table 17) between SREE1 (*P. merguiensis*) and NKE8 (*P. silasi*), at the third column and the fourth row, was about 0.0054 (0.54%), between SREE1 (*P. merguiensis*) and TA5 (*P. indicus*), the third column and the third row, about 0.0145 (1.45%), and between SREE1 (*P. merguiensis*) and D1 (*P. monodon*), at the third column and the fifth row, about 0.0616 (6.16%). There was a low level of genetic distance between SREE1, NKE8 and TA5. Matthews et al. (2002) used that the nucleotide divergence of 16S rRNA 1-3% to show the relationships between two snapping prawn species (Crustacea: Decapod *Alpheus*). Therefore the distance between STWA6 and TA5 (about 2.57%) seemed high for genetic divergence within species.

The genetic distance between SREE1 and D1 was the highest distance. This result corresponded with morphological data that *P. monodon* is clearly identified from other species.

The neighbor-joining tree (Figure 20) from the nucleotide sequence indicates that STWA6 is closer to SREE1 (*P. merguiensis*) than TA5 (*P. indicus*).

```
                *        20         *        40         *        60
SREE1 : CCGGTCTGAACTCAGATCACGTTTGGATTTAAAGGTCGAACAGACCTTGCTTTATAACTG :  60
NKE8  : ......................AA.........................C........... :  60
STWA6 : ......................AA.........................C........... :  60
TA5   : .........-............AA...................C....C........... :  59
D1    : ..............A....T..AA.........................C....G...... :  60

                *        80         *       100         *       120
SREE1 : CTGCATCATAAGGATACCTTAATTCAACATCGAGGTCGCAAACCTTCTTGTCGATAGGGA : 120
NKE8  : ..................................................................... : 120
STWA6 : .....................................................T... : 120
TA5   : ..................................................................... : 119
D1    : ......T.....................................................GA... : 120

                *       140         *       160         *       180
SREE1 : CTCTCAAAGAAGATTACGCTGTTATCCCTAAAGTAACTTAATCTTTTAATCGCTAATAGA : 180
NKE8  : ..................................................................... : 180
STWA6 : ..................................................................... : 180
TA5   : ..................................................................... : 179
D1    : ..................................................TT.....A. : 180

                *       200         *       220         *       240
SREE1 : GGATCATACTAATTTTCAATTATATTTGTTATTAAA-TATTTAAGAACAGTTACCTATTA : 239
NKE8  : ......C.............................-...................... : 239
STWA6 : ......C..................C.........-...................... : 239
TA5   : ..............C......................-...................... : 238
D1    : ......-.T..T.C..........C......A....A.................TA.... : 239

                *       260         *       280       ●  *       300
SREE1 : TATTCTCGTCGCCCCAACGCAACAAATTTTAATTAAAACCAAGCTACACTAACAATTGAT : 299
NKE8  : ............................C.............T..T............. : 299
STWA6 : ..........................................T.............. : 299
TA5   : ............................A.............T..T............. : 298
D1    : .....C.....................CA..........T....T..T..........T.. : 299
```

Figure 19. The nucleotide sequence alignment of 16S rRNA gene in SREE1, NKE8, STWA6, TA5 and D1. ● Indicate the informative site. - Indicate the deletion position.

Figure 19. (continued)

```
                *       320         *       340         *       360
SREE1 : AGTATAATTAAATTATTGTCAAGCTTTATAGGGTCTTATCGTCCCCCTAATTTATTTAAG : 359
NKE8  : ..............................................T.............. : 359
STWA6 : ..............................................T.............. : 359
TA5   : ..............................................T.............. : 358
D1    : .A.T....C..........T...T.....................T...AG........ : 359

                *       380         *       400         *       420
SREE1 : CCTTTTCACTTAAAAGTTAAGTTCAATTATTATAATTGAGACAGCTTACTTTTTGTCCAA : 419
NKE8  : ............................C................................ : 419
STWA6 : ..................................................................... : 419
TA5   : ......................A.......................................... : 418
D1    : ............................C........C........A................ : 419

                *       440         *       460         *       480
SREE1 : CCATTCATACAAGCCTTCAATTAAAAGACTAATGATTATGCTACCTTCGCACGGTCAATA : 479
NKE8  : ..................................................................... : 479
STWA6 : ..................................................................... : 479
TA5   : ..................................................................... : 478
D1    : .....................................................G.. : 479

                *       500         *       520         *       540
SREE1 : TACCGCGGCCCTTTAAAC-TAAATCAGTGGGCAGGCTAGACTTTATATAACAATCATATA : 538
NKE8  : ..................-.......................................... : 538
STWA6 : ..................-.....................................A.A... : 538
TA5   : ..................-.....................................C...... : 537
D1    : ..................AA...T..................................... : 539

                *       560
SREE1 : GACATGTTTTTGTTAAACAGGCG : 561
NKE8  : ............G.......... : 561
STWA6 : ....................... : 561
TA5   : ....................... : 560
D1    : ............A.........A : 562
```

Table 16. The nucleotides occurring at variable positions in the 16S rRNA mtDNA genes of STWA6, SREE1 (*P. merguiensis*), NKE8 (*P. silasi*), TA5 (*P. indicus*) and D1 (*P. monodon*).

Positions	STWA6	SREE1	NKE8	TA5	D1
15	G	G	G	**G**	**A**
20	C	C	C	C	**T**
44	A	A	A	**C**	A
54	A	A	A	A	**G**
67	C	C	C	C	**T**
116	A	A	A	A	**G**
117	**T**	G	G	G	**A**
172-173	GC	GC	GC	GC	**TT**
179	G	G	G	G	**A**
189	C	C	C	C	**T**
192	A	A	A	A	**T**
194	T	T	T	T	**C**
196	T	T	T	**C**	T
205-206	**TC**	TT	TT	TT	**CT**
212	T	T	T	T	**A**
235-236	CT	CT	CT	CT	**TA**
246	T	T	T	T	**C**
267-268	**TT**	**TT**	**TC**	**TA**	**CA**
279	C	C	C	C	**T**
284, 287	**T, C**	**C, C**	T, T	T, T	T, T
298	G	G	G	G	**T**
302	G	G	G	G	**A**
304	A	A	A	A	**T**
309	T	T	T	T	**C**
320	C	C	C	C	**T**
324	C	C	C	C	**T**
351-352	TT	TT	TT	TT	**AG**
381	G	G	G	**A**	G
387-388	TT	TT	**TC**	TT	**CT**
405	C	C	C	C	**A**
478	A	A	A	A	**G**
503	A	A	A	A	**T**
535	**A**	T	T	**C**	T
554	T	T	**G**	T	T
562	A	A	A	A	**G**

Table 17. The distance matrix of the 16S sequences of STWA6, SREE1 (*P. merguiensis*), NKE8 (*P. silasi*), TA5 (*P. indicus*) and D1 (*P. monodon*).

	STWA6*	SREE1	TA5	NKE8	D1
STWA6	0.0000				
SREE1	0.0145	0.0000			
TA5	0.0257	0.0145	0.0000		
NKE8	0.0163	0.0054	0.0145	0.0000	
D1	0.0718	0.0616	0.0696	0.0596	0.000

*COI sequence was classified as *P. merguiensis*

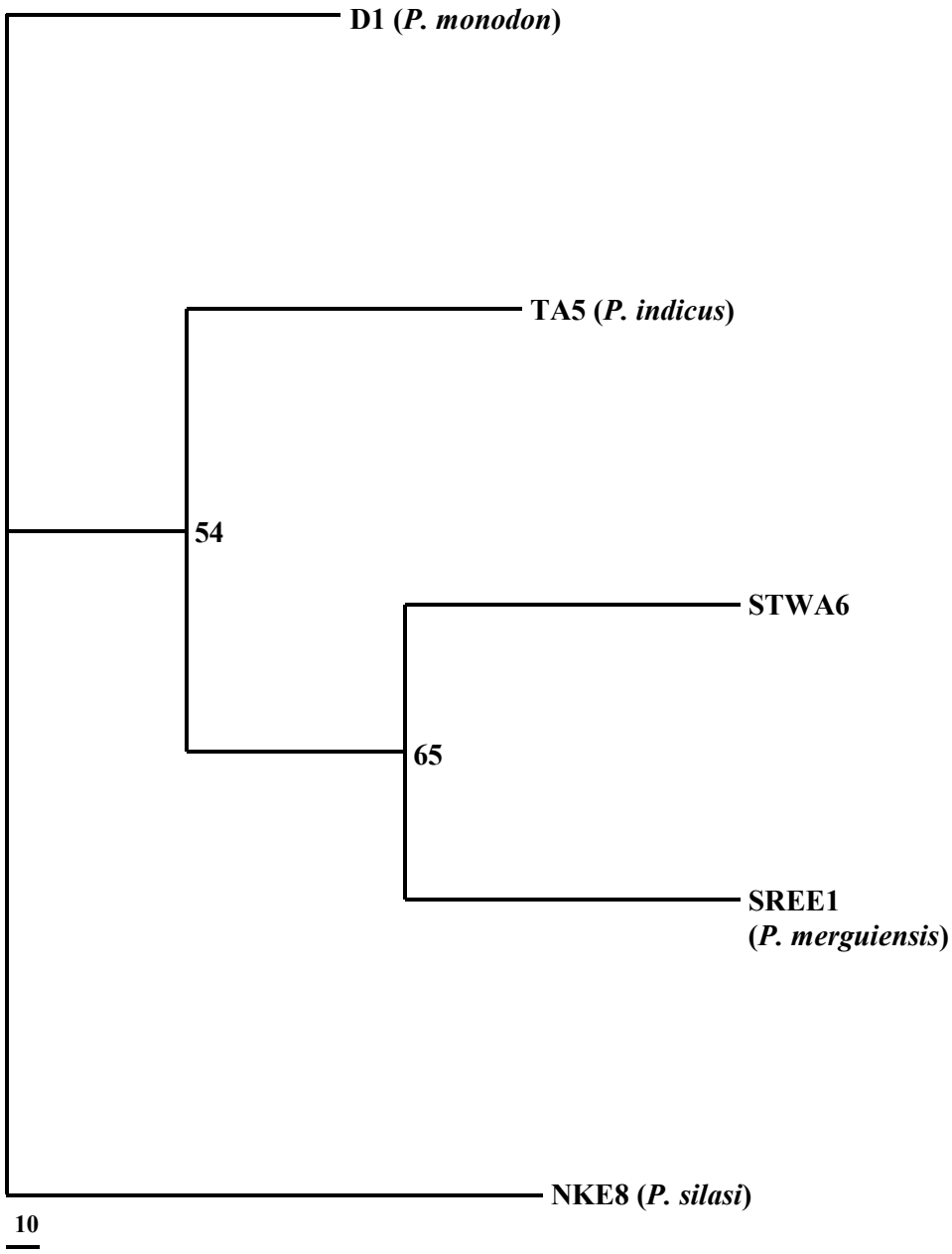

Figure 20. The neighbor-joining tree of 16S nucleotide sequences of sample STWA6, SREE1, NKE8, TA5 and D1. — Indicates the scaled branches phylogenetic tree: lengths of branches are proportional to the number of molecular changes.

The results from 12S and 16S data indicated that STWA6 was not likely to be the same species as *P. indicus* because of the large difference in genetic distances between STWA6 and TA5 from the 12S data (7.78%) and also in the 16S data (2.5%). The data showed too much divergence within the same species (*P. indicus*). When compared with the genetic distance of 16S data between *P. merguiensis* and *P silasi*, it was only 0.54% and between *P. merguiensis* and *P. indicus* was only 1.45%.

The neighbor-joining trees of 12S- and 16S-nucleotide sequences are shown (Figure 18 and 20). The bootstrap value supported that STWA6 clustered in *P. merguiensis* (SREE1). It seemed that STWA6 was not *P. indicus* because it was not clustered in *P. indicus* (TA5). We suspected that STWA6 could be a genetic variation of *P. merguiensis* and the identification by morphology is misleading.

Because of the suspicion above, the COI gene (mitochondrial protein coding) sequence was chosen for confirmation.

3.3. Cytochrome oxidase subunit I (COI) gene

Referring to the discussion above concerning 12S and 16S data, here we confirmed the results through COI analysis. Cytochrome Oxidase is present in all organisms having mitochondrial respiratory chains: plants, animals, and other eukaryotic organisms. Its function and therefore sequence is highly conserved. This makes logical sense in that mutations occurring in its sequence could be fatal for the dynamics of cellular respiration. The cytochrome oxidase I gene (COI) has two important advantages. Firstly, the universal primers for this gene are very robust, enabling recovery of its 5' end from the representatives of most, if not all, animal

phyla (Hebert et al., 2003). Secondly, COI likely possesses a greater range of phylogenetic signals than any other mitochondrial gene. In common with other protein-coding genes, its third-position nucleotides show a high incidence of base substitutions. However, changes in its amino acid sequence occur more slowly than those in other mitochondrial genes.

Baldwin et al. (1998) found the great divergence of 558 bp of nucleotide sequence in the COI gene and it was a good molecular marker for phylogenetic study among 13 species representing all six subgenera of the prawn genus *Penaeus*. Analyses of this sequence revealed high genetic divergence between species (8-24%). The nucleotide divergence ranged from 0-3% within species and 11-19% within subgenera.

Because the COI nucleotide sequence provided a high level of divergence between species and was also easily and clearly differentiated, COI was the next target gene for our research study in *Penaeus* (*Fenneropenaeus*)

To begin, the thirty-two samples were chosen according to the following criteria:

Based on the morphological character of each species (*P. merguiensis*, *P. silasi* and *P. indicus*). In theoretical, if there are three groups of sequences and these come from three animals species that are morphologically different (with the difference being *P. merguiensis*, *P. silasi* and *P. indicus*) then we could look at differences of mitochondrial DNA genotype within each of these three species. Our study, thirty-two specimens from eight locations was selected.

Five specimens (NKE2, NKE5, NKE8, NKE17 and NKE19 from Nakhon si thummarat) of certain *P. silasi* species were selected based on morphological characteristics. *P. merguiensis* is the major species of white prawn in Thailand. Then three samples, STWA5, STWA6 and STWA7 were strongly identified as *P. indicus*, and two samples, SREE1 and SREE2, were strongly identified as *P. merguiensis* were chosen. Twenty-two specimens came from groups I, II and III by morphological method were also chosen. Chaitiamvong and Supongpan (1992) did not found *P. indicus* in Thai waters, but we needed *P. indicus* as reference to compare with the sequence data of STWA6. So we contacted a laboratory in South Africa (where the major species of white prawn is *P. indicus*) and obtained the DNA of *P. indicus* from there to use as a reference. If STWA6 was *P. indicus,* the results should be the same as the reference. Four specimens of *P. indicus* species (PIZA, PIOMAN, PIMZ and PITA5) were chosen. In addition, construction of a tree may be aided by getting outgroup sequences (*P. monodon* and *P. vannamei*) as well. *P. indicus* and *P. merguiensis* sequences in database is required for cluster same species together and separate from different species.

Table 18. The thirty-two samples chosen for phylogenetic study.

Samples name	Note
STWA5, STWA6 and STWA7 (*P. indicus*) SREE1, SREE2 (*P. merguiensis*)	Old samples that were positively identified through thelycum features
PKW7, SKE26, SKEB1, SKEB5, TDE1, TDE23, RNW55, STW4, STW31, TRW7, NKE1, NKE9	Represent group I by morphological and morphometric identification
SKE43, SKE92, TRW11, TRW15, NKE18	Represent group II by morphological and morphometric identification
SKE96, TDE2, TDE12, SRE30, RNW66	Represent group III by morphological and morphometric identification
NKE2, NKE5, NKE8, NKE17 and NKE19	Represent as *P. silasi* by morphological and morphometric identification
PIMZ, PITA5, PIZA and PIOMAN	Represent as *P. indicus*
PIGB (*P. indicus*) PMERG (*P. merguiensis*) PMOGB (*P. monodon*) PVANGB (*P. vannamei*)	Sequences retrieved from the GenBank

The amplification of the 5'end of the COI gene with primer mtD-4 and mtD-9 is shown in Figure 21, lanes 2 to 4. This fragment was about 800 bp in length. The 3'end fragment of COI amplified with primer mtD-8 and mtD-12 is shown in lanes 5 to 7 (Figure 21). This fragment was about 900 bp in length.

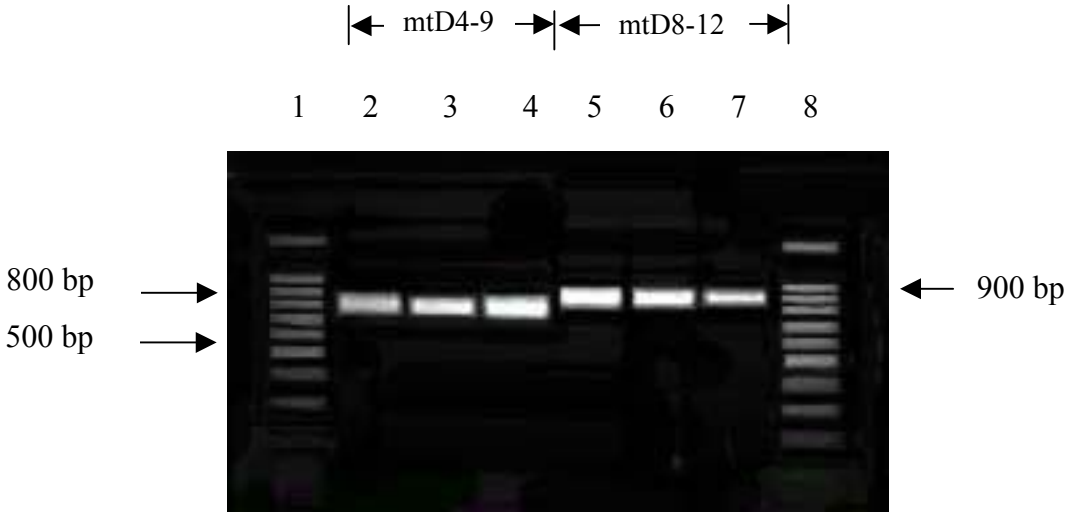

Figure 21. The PCR product of about 800 bp of COI gene amplified by using primers mtD-4 and mtD-9 (lane 2-4), and primers mtD-8 and mtD-12 (lane 5-7). The 100-bp markers are shown in lanes 1 and 8.

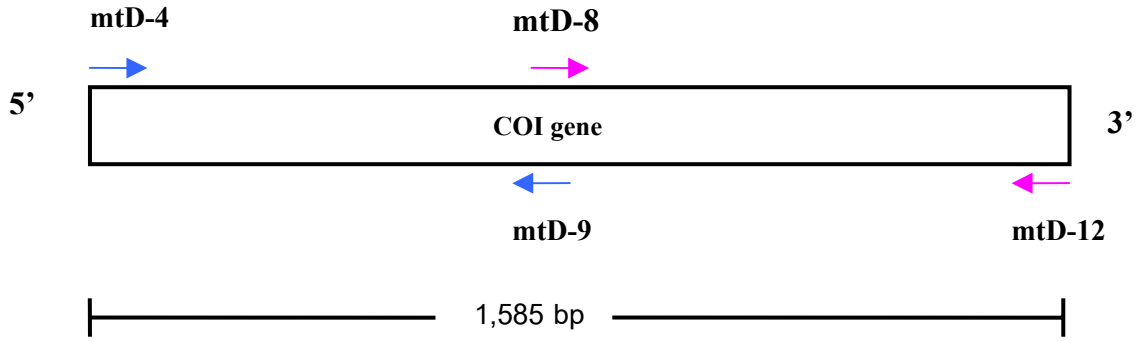

Figure 22. The complete COI nucleotide sequence from using two overlapping primer pairs.

The obtained sequences of SREE1 were compared with the BLASTN program in the GenBank database (www.ncbi.nih.nlm.gov/BLAST). They showed 99% identity with Cytochrome oxidase subunit I (COI) gene of *P. merguiensis* from Singapore (Accession number: AY143988.1, e-value = 0.0), 96% identity with COI gene of *P. merguiensis* from Australia (Accession number: AY143986.1, e-value = 0.0), 95% identity with COI gene of *P. merguiensis* (Accession number: AF029390.1, e-value = 0.0). The complete nucleotide of this gene was done by sequence overlapping between primer mtD-8 and mtD-9 (Figure 22). The complete nucleotide sequence of this *P. merguiensis* is composed of 1,585 bp in length (Figure 23).

TACAATTTAT CGCCTAAACT TCAGCCATTT TATCTCAACG CAACGATGATTAT

TTTCTACAAA TCATAAAGAC ATTGGAACCT TATACTTTAT CTTCGGAGCC

TGAGCTGGAA TAGTAGGGAC TGCCCTTAGA CTTATTATTC GTGCTGAATT

GGGTCAACCG GGAAGCCTCA TTGGAGATGA TCAAATTTAT AATGTGGTTG

TTACAGCCCA CGCTTTCGTT ATAATTTTCT TCATGGTAAT ACCTATTATA

ATTGGAGGAT TTGGAAATTG ACTAGTCCCT TTAATATTAG GTGCTCCCGA

TATAGCTTTC CCTCGAATAA ATAATATGAG TTTCTGACTT TTACCTCCCT

CGCTAACTCT ACTTCTTTCT AGAGGTATAG TCGAAAGAGG GGTAGGAACA

GGATGAACTG TCTACCCTCC TCTATCTGCC AGCATTGCCC ATGCAGGAGC

ATCTGTAGAT CTAGGAATCT TCTCATTACA TTTAGCAGGG GTTTCTTCAA

TTCTAGGTGC TGTAAATTTT ATAACAACTG TTATCAATAT ACGATCAACA

GGAATAACTA TAGACCGAAT ACCTCTTTTC GTCTGAGCGG TATTTATTAC

AGCCTTACTA CTTTTACTAT CATTACCAGT TTTAGCGGGA GCTATTACAA

TGCTTCTAAC GGACCGAAAC CTAAATACTT CATTCTTCGA TCCTGCAGGA

GGGGGAGATC CTGTTTTATA T**CAACACTTA TTTTGATTCT TCGG**TCACCC

A***GAAGTTTAT ATTTTAATTT TACC***TGCCTT CGGAATGATT TCACACATTA

TTAGCCAAGA ATCAGGTAAG AAAGAAGCAT TTGGAACTCT CGGAATAATT

TATGCTATAC TGGCTATTGG TGTACTAGGA TTTGTAGTTT GAGCACACCA

TATATTTACA GTAGGTATGG ATGTTGATAC TCGTGCTTAC TTTACATCTG

CCACAATAAT TATTGCTGTT CCTACGGGCA TCAAAATTTT TAGCTGACTA

GGAACACTTC ACGGTACTCA GCTAAACTAT AGCCCCTCTC TAATCTGAGC

ACTAGGATTT GTATTTTTAT TTACTGTAGG AGGCCTTACA GGAGTAGTTC

TTGCCAACTC ATCAATTGAT ATTATTTTAC ATGATACATA TTATGTTGTT

GCCCATTTCC ATTATGTTCT TTCTATAGGA GCAGTATTTG GTATTTTTGC

TGGTATTGCC CACTGATTCC CACTCTTTAC AGGACTTACT TTAAACCCCA

AATGACTAAA AATCCACTTT ATAGTAATGT TTATTGGAGT AAATATTACA

TTCTTCCCTC AACATTTCTT AGGACTTAAT GGAATACCTC GGCGGTACTC

TGACTACCCA GACGCCTATT CAGCATGAAA TGTCGTTTCA TCTATTGGAT

CTACAGTATC ATTGATTGCC GTTTTAGGAT TTGTTATAAT TGTATGAGAA

GCATTAACCG TTGCTCGACC AGTTATATTC TCTCTCTTCC TACCAACTTC

AATTGAATGA CAACATAATC TTCCGCCTGC AGACCACAGT TATATAGAAA

ATCCTTTAAT TACTAACTTC **TAATATGGCA GATTAGTGCA TTGGA**

Figure 23. The complete nucleotide sequence of COI (Accession No. AF284432). The bold and italic letters show the mtD-4 and mtD-9 primers, the bold letters show the mtD-8 and mtD-12 primers. The mtD-9 primer overlaps the mtD-8 primer. The underlined letter is the 558-bp nucleotide region used for phylogenetic construction.

3.3.1 558 bp analysis at the 3'end of COI gene

From sequencing, the 3' end of the COI gene (874 bp), we compared the 558 bp region among the samples mentioned above (Table 18). Baldwin et al. (1998) reported that this region showed high genetic divergence between species in *Penaeus sp*. Gusmão et al. (2000) used that region to discriminate among the four main Brazilian penaeid species and detected a new species of *Penaeus*.

The percentage of AT and GC composition in the 558 bp region is about 63% and 37%, respectively. From the sequence alignment, there were 178

substitution sites and 28 informative sites. Transitions were more numerous than transversions (111:34). There was no length variation among the 558 bp of COI sequences, corresponding to many reports (i.e. Baldwin et al., 1998; Gusmão et al., 2000). Transitions from C-T were also more commonly found than A-G, similar to 12S and 16S data. The A-T transversions predominated over other transversions, as in insect mitochondrial DNA (Simon et al., 1994).

3.3.2 Two distinct clusters of *P. merguiensis* based on COI sequence

A neighbor-joining tree (Figure 24), constructed based on sequence divergence between pairs of sequences, separated specimens into four clusters; A, B, C and D. *P. merguiensis* formed two distinct clusters A and B. The topology was a paraphyletic group. Cluster A mainly consisted of *P. merguiensis* from the Gulf of Thailand, and also included *P. merguiensis* (GB) from Taiwan in the Pacific Ocean. However, sample STWA5, STWA6 and STWA7, strongly identified as *P. indicus*, were placed in cluster A. Sample SREE1 and SREE2, strongly identified as *P. merguiensis*, were placed in cluster B, which mainly included *P. merguiensis* from the Andaman Sea (Table 19). Cluster C was composed of the samples from Nakhon Si Thummarat, which were clearly identified by using the third maxilliped in male as *P. silasi*, including the female sample, NKE18, which could not be determined as to species clearly by morphology alone. Cade D contained the *P. indicus* samples from South Africa, grouped in the same, section as *P. indicus* from the GenBank database.

The distances between clusters A and B showed an average 5.0% divergence. The distance between *P. merguiensis* and cluster C, *P. silasi* was 8.61%

and the distance between *P. merguiensis* and *P. indicus,* cluster D is 11.98% (Table 20). From the results, these species could be distinguished clearly, and the relationship between *P. merguiensis* and *P. silasi* seen to be closer than *P. indicus*.

The 186 amino acids of this region were aligned (Figure 25) by Clustal W and the Neighbor-Joining tree constructed (Figure 26). The distance of the amino acid sequences between *P. indicus* and other sequences were between 0.00557-0.3386, between *P. monodon* and other sequences 0-0.02314 (Table 21), and between *P. indicus* and *P. merguiensis* was 0.03377. In addition the distance of the amino acid sequences between *P. monodon* and *P. vannamei* were not different. The neighbor-joining tree showed a weak bootstrap value (<50) of most specimens and could not cluster them as did the morphological data. These results indicate that the amino acid sequences of COI are very conserved in closely related species and can not be used to separate each species clearly. Therefore the amino acid sequences are not appropriate for phylogentic study in closely related species (low level of genetic distance).

3.3.3 Genetic variation within cluster

COI sequence variation within species ranged from 1-2% (Bucklin et al., 1998). In 2000, Gusmão et al. found the highest level of intraspecific COI sequence divergence was 1.26% with *Penaeus subtilis*.

In our results, there was substantial COI variation within cluster A and cluster B of *P. merguiensis*. The divergence within clusters A and B are 0.0-2.25% and 0.18-1.49%, respectively. However, this variation did not show strong geographical and morphological structuring within clusters. For example, TDE

individuals collected on the same day from the same locality were in different parts of the clusters. TDE1 showed morphological and locality difference from STWA5, but were closely aligned in cluster

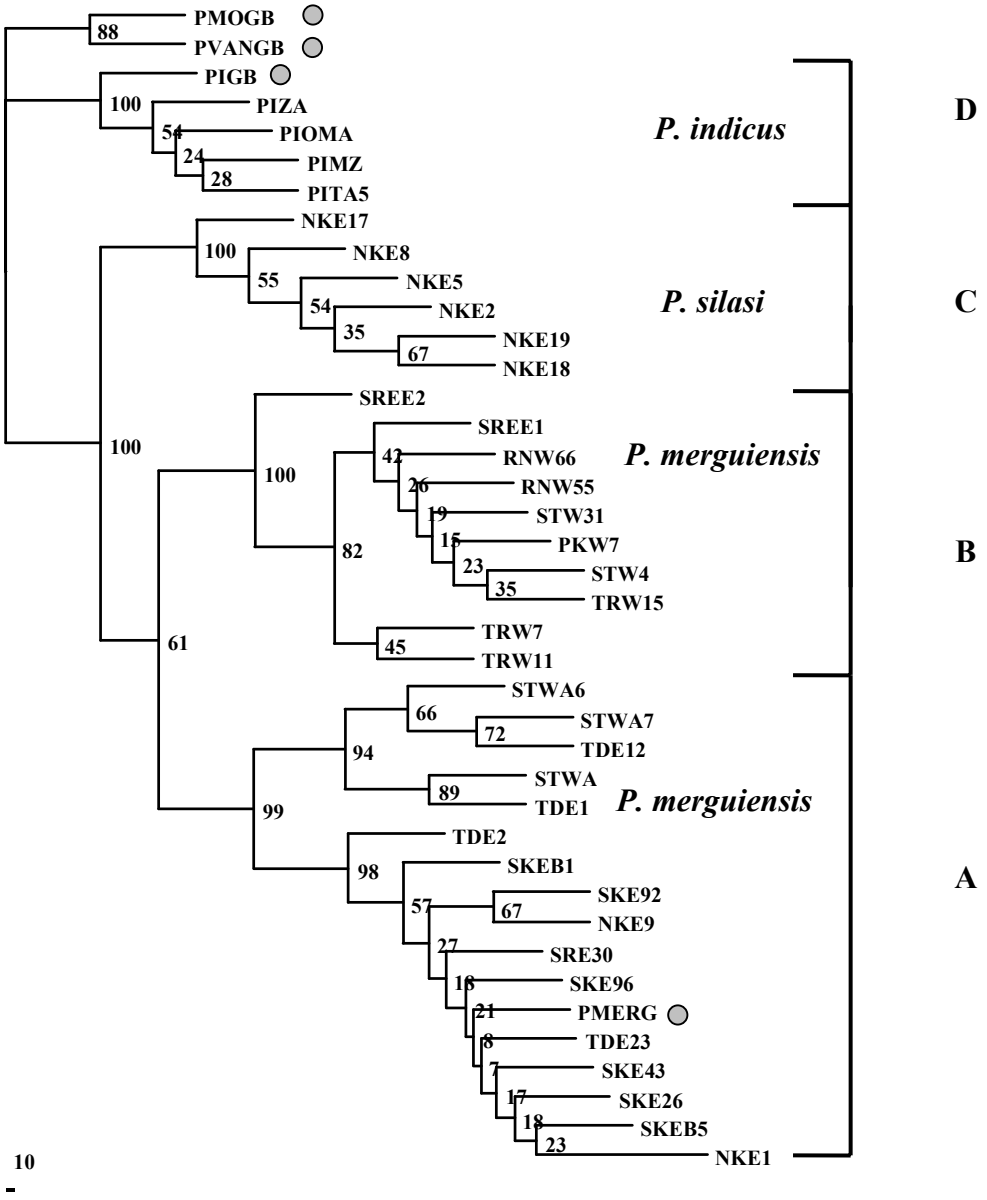

Figure 24. **A neighbor-joining tree constructed based on the percentage of sequence divergence between pairs of COI sequences. The percent bootstrapping values (1000 replicates) between branching groups are indicated. Sequence reference from GenBank (○) (*P. merguiensis* (GB); *P.indicus* (GB); *P. monodon* (GB); *P. vannamei* (GB)). — Indicates the scaled branches phylogenetic tree: lengths of branches are proportional to the number of molecular changes.**

Table 19

Table 20

```
                 *        20         *        40         *        60
PIZA    : MISHIISQESGKKEAFGTLGMIYAMLAIGVLGFVVWAHHMFTVGMDVDTRAYFTSATMII :  60
PIOMAN  : ............................................................ :  60
PITA5   : ............................................................ :  60
PIMZ    : ............................................................ :  60
PIGB    : ............................................................ :  60
NKE18   : ............................................................ :  60
NKE19   : ............................................................ :  60
NKE2    : ............................................................ :  60
NKE5    : ............................................................ :  60
NKE8    : ............................................................ :  60
NKE17   : ............................................................ :  60
NKE1    : ............................................................ :  60
SKEB5   : ..............................S............................. :  60
SKE96   : ............................................................ :  60
PMERG   : ............................................................ :  60
TDE23   : ............................................................ :  60
SKE26   : ............................................................ :  60
SKE43   : ............................................................ :  60
NKE9    : ............................................................ :  60
SKE92   : ............................................................ :  60
SKEB1   : ............................................................ :  60
SRE30   : ............................................................ :  60
TDE2    : ............................................................ :  60
TDE1    : ............................................................ :  60
STWA5   : ............................................................ :  60
TDE12   : ............................................................ :  60
STWA7   : ............................................................ :  60
STWA6   : ............................................................ :  60
PKW7    : ............................................................ :  60
STW31   : ...........................................-................ :  59
TRW15   : ............................................................ :  60
STW4    : ............................D............................... :  60
RNW66   : ............................................................ :  60
SRE1    : ............................................................ :  60
RNW55   : ...................S........................................ :  60
TRW11   : ............................................................ :  60
TRW7    : ............................................................ :  60
SRE2    : ....................................L....................... :  60
PMOGB   : ............................................................ :  60
PVANGB  : ............................................................ :  60
```

Figure 25. The multiple alignment of the amino acid sequences of COI gene (from 558 bp nucleotide sequence) with that of various COI amino acid sequences; *P. indicus* (PIGB), *P. merguiensis* (PMERG), *P. monodon* (PMOGB) and *P. vannamei* (PVANGB)

Figure 25. (continued)

```
                 *        80         *       100         *       120
PIZA    : AVPTGIKIFSWLGTLHGTQLNYSPSLIWALGFVFLFTVGGLTGVVLANSSIDIILHDTYY : 120
PIOMAN  : ................D.......................................... : 120
PITA5   : ...............R..............S............................ : 120
PIMZ    : ..................................................................... : 120
PIGB    : ..................................................................... : 120
NKE18   : ..................................................................... : 120
NKE19   : ...................................................T..... : 120
NKE2    : ..................................................................... : 120
NKE5    : ..................................................................... : 120
NKE8    : ..................................................................... : 120
NKE17   : .....................................................................C : 120
NKE1    : ..................................................................... : 120
SKEB5   : ..................................................................... : 120
SKE96   : ..................................................................... : 120
PMERG   : ..................................................................... : 120
TDE23   : ..................................................................... : 120
SKE26   : ...................................................T..... : 120
SKE43   : ..................................................................... : 120
NKE9    : ..................................................................... : 120
SKE92   : ..................................................................... : 120
SKEB1   : ..........................................S.......... : 120
SRE30   : ........L..................................................... : 120
TDE2    : ..................................................................... : 120
TDE1    : ..................................................................... : 120
STWA5   : ..................................................................... : 120
TDE12   : ..................................................................... : 120
STWA7   : ..................................................................... : 120
STWA6   : ..................................................................... : 120
PKW7    : ..................................................................... : 120
STW31   : ..................................................................... : 119
TRW15   : ..................................................................... : 120
STW4    : ........................P.............................-. : 119
RNW66   : ..........................................E.......... : 120
SRE1    : ..................................................................... : 120
RNW55   : ..................................................................... : 120
TRW11   : ..................................................................... : 120
TRW7    : ..................................................................... : 120
SRE2    : ..................................................................... : 120
PMOGB   : ..................................................................... : 120
PVANGB  : ..................................................................... : 120
```

Figure 25. (continued)

```
                  *       140         *       160         *       180
PIZA   : VVAHFHYVLSMGAVFGIFAGIAHWFPLFTGLTLNPKWLKVHFLVMFIGVNITFFPQHFLG : 180
PIOMAN : ..................................................................... : 180
PITA5  : ..................................................................... : 180
PIMZ   : ......C.............................................................. : 180
PIGB   : .....................................------------------------------- : 163
NKE18  : ..................................................................... : 180
NKE19  : .....................................................V............... : 180
NKE2   : ..................................................................... : 180
NKE5   : ..................................................................... : 180
NKE8   : ..................................................................... : 180
NKE17  : ..................................................................... : 180
NKE1   : ..................................................................... : 180
SKEB5  : ..................................................................... : 180
SKE96  : ..................................................................... : 180
PMERG  : ..................................................................... : 180
TDE23  : ..................................................................... : 180
SKE26  : ..................................................L.................. : 180
SKE43  : ..................................................................-. : 179
NKE9   : ..................................................................S : 180
SKE92  : ..................................................................... : 180
SKEB1  : ..................................................................... : 180
SRE30  : ..................................................T.................. : 180
TDE2   : ..............................-...................................... : 179
TDE1   : .............................................V..G.................... : 180
STWA5  : .........P.....................................V.............. : 180
TDE12  : ..................................................................... : 180
STWA7  : ..................................................................... : 180
STWA6  : ..................................................................... : 180
PKW7   : ..................................................................... : 180
STW31  : ..................................................................... : 179
TRW15  : ....S................................................................ : 180
STW4   : ..................................................................... : 180
RNW66  : ..................................................................... : 180
SRE1   : ..................................................................... : 180
RNW55  : .....................................................V.........L.... : 180
TRW11  : ..................................................................S : 180
TRW7   : ..................................................................S : 180
SRE2   : ..................................................................... : 180
PMOGB  : ..................................................................... : 180
PVANGB : ..................................................................... : 180
```

Figure 25. (continued)

```
PIZA   : LNGMPR : 186
PIOMAN : ...... : 186
PITA5  : ...... : 186
PIMZ   : ...... : 186
PIGB   : ------ :   -
NKE18  : ...... : 186
NKE19  : ...... : 186
NKE2   : ...... : 186
NKE5   : ...... : 186
NKE8   : ...... : 186
NKE17  : ...... : 186
NKE1   : ...... : 186
SKEB5  : ...... : 186
SKE96  : ...... : 186
PMERG  : ...... : 186
TDE23  : ...... : 186
SKE26  : ...... : 186
SKE43  : ...... : 185
NKE9   : ...... : 186
SKE92  : ...... : 186
SKEB1  : ...... : 186
SRE30  : ...... : 186
TDE2   : ...... : 185
TDE1   : ...... : 186
STWA5  : ...... : 186
TDE12  : ...... : 186
STWA7  : ...... : 186
STWA6  : ...... : 186
PKW7   : ...... : 186
STW31  : ...... : 185
TRW15  : ...... : 186
STW4   : ...... : 185
RNW66  : ...... : 186
SRE1   : ...... : 186
RNW55  : ...... : 186
TRW11  : ...... : 186
TRW7   : ...... : 186
SRE2   : ...... : 186
PMOGB  : ...... : 186
PVANGB : ...... : 186
```

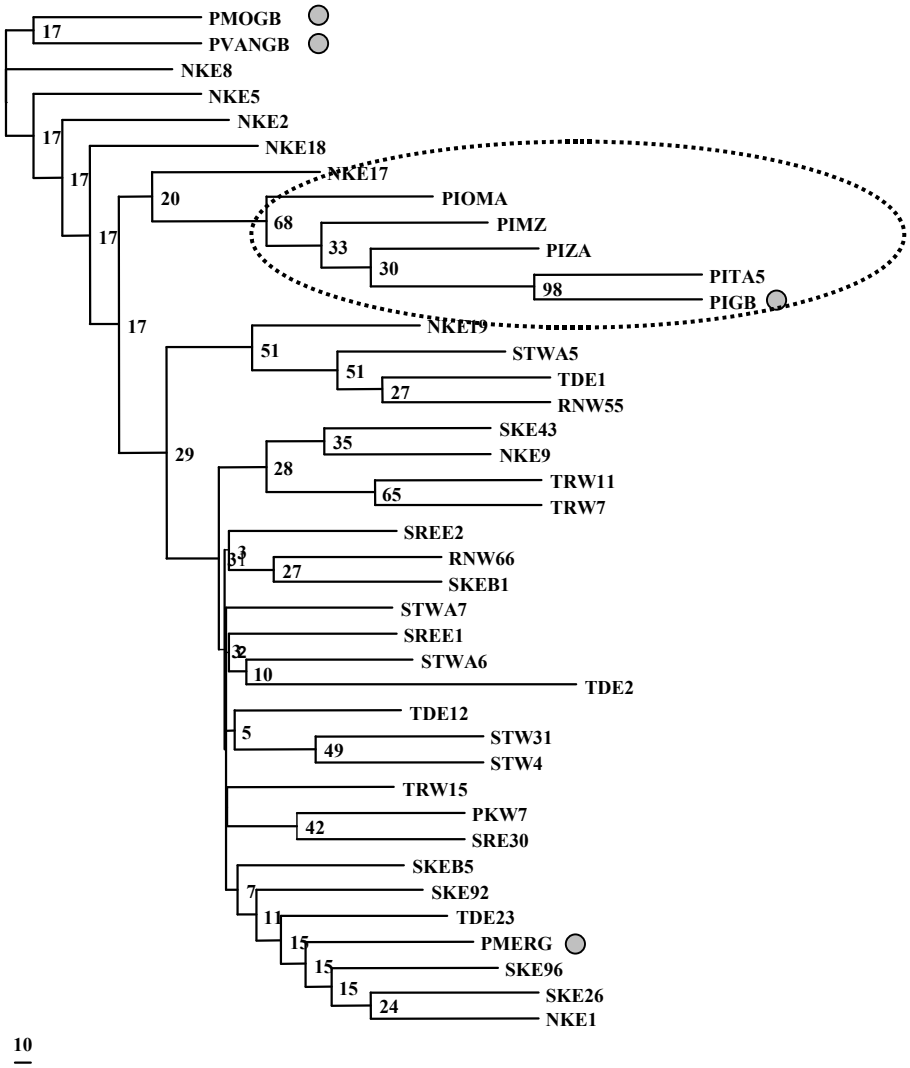

Figure 26. A neighbor-joining tree constructed based on amino acid sequence divergence between pairs of COI sequences. The percent bootstrapping values (1000 replicates) between branching groups are indicated. Sequence reference from GenBank (○)(*P. merguiensis* (GB); *P.indicus* (GB); *P. monodon* (GB); *P. vannamei* (GB)).

— Indicates the scaled branches phylogenetic tree: lengths of branches are proportional to the number of molecular changes.

Table 21. distance amino acid

3.4 PCR-RFLP analysis of the 5'end and 3' end of COI gene

PCR-RFLP, direct sequencing and amplification-based profiling techniques are useful for species analysis and phylogenetic reconstruction. Gusmão et al. (2000) used a restriction fragment length polymorphism (RFLP) analysis of the COI gene to obtain species-specific markers for the identification of larvae and industrial products of *Penaeus* (*Farfantepenaeus*).

Based on the sequence data of the COI gene, a computer search was performed using the program DNASIS (version 2.0) for one or more restriction endonucleases that could produce diagnostic COI fragment profiles among closely related target species. All six taxa (*P. merguiensis* (SREE1, N=1), *P. silasi* (NKE8, N=1), *P. indicus* (TA5, N=1), *P. monodon* (D1, N=1), *P. semisulcatus* (from Songkhla, N=1) and *Metapenaeus* (from Songkhla, N=1)) produced the expected polymerase chain reaction (PCR) product of approximately 900 bp (Figure 27A). The four-base recognition endonuclease *Rsa*I was potentially informative in the 3' end region of COI gene. Digestion with *Rsa*I produced a diagnostic fragment pattern for each of the closely related target taxa (*P. merguiensis*, *P. silasi* and *P. indicus*) (Figure 27B). Table 22 shows that the digestion of the 3' end of COI produced three fragments in *P. merguiensis*, two fragments (300, 582) in *P. silasi* and three fragments (282, 12, 248, 334) in *P. indicus*. Two fragments (295, 555) were produced from *P. monodon* and *P. semisulcatus* (376, 524), and they differed between recognized species. A search sequence in *P. merguiensis* indicated four bands (142, 152, 240, 360) are expected, but there were only three bands (152, 240, 360) on the gel. The summed fragment length for each taxon was generally closed to the expected 900 bp,

except for *P. merguiensis*, *P. indicus* and *Metapenaeus sp* because some small fragments of 20-50 bp could not be visualized on the agarose gel stained with Ethidium Bromide.

From the sequencing of the 5' end of the COI gene, this fragment was cleaved by restriction enzymes (*Mbo*I and *Bgl*II) and given different profiles between the *P. merguiensis* specimens from the Gulf of Thailand and the Andaman Sea (East and West coasts of Thailand, respectively) (Figure 28). The specimens from the Gulf of Thailand (Trad, Nakhon Si Thummarat, Surat Thani and Songkhla) gave the haplotype 2, 4 and the fragment sizes are shown in Table 23. The specimens from the Andaman Sea (Ranong, Trang and Satun) gave the haplotype 1, 3 and the fragment sizes are also shown in Table 23. Using the PCR-RFLP technique, the specimens of *P. merguiensis* were mostly classified by the geographic origin from the East and West coasts of Thailand, which also corresponded to the 3' end sequence analysis in samples from cluster A (the East) and cluster B (the West).

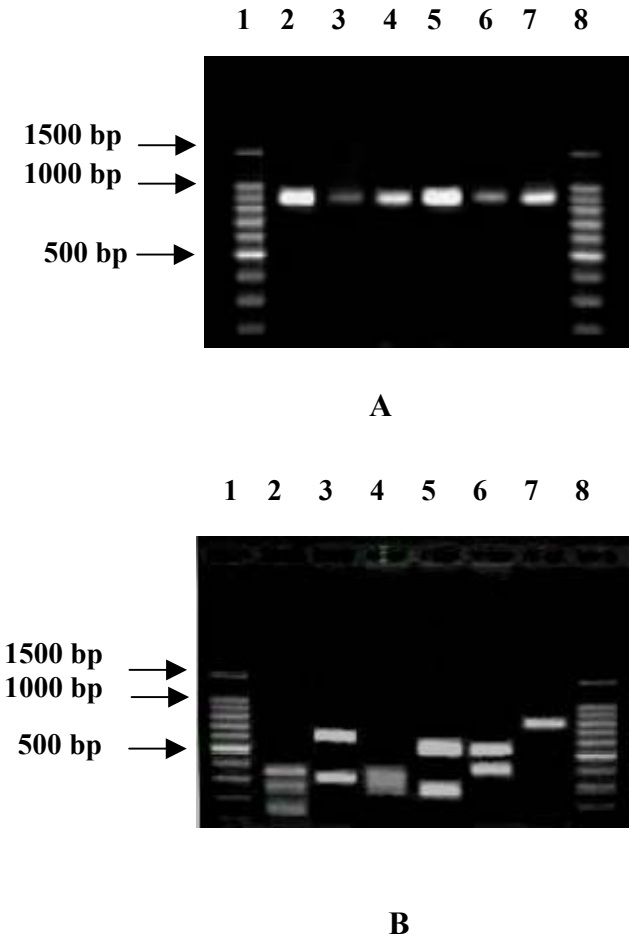

Figure 27. A) The PCR products of 3'end of the COI fragment (primer mtD-8, mtD12). B) The restriction pattern of the 3' end of the COI cleaved by *Rsa*I. Lanes 1 and 8 = 100 bp marker; lane 2 = *P. merguiensis*, lane 3 = *P. silasi*, lane 4 = *P. indicus,* lane 5 = *P. monodon,* lane 6 = *P. semisulcatus* and lane 7 = *Metapenaeus*.

Table 22. The amplified fragments of COI (primer mtD-8 and mtD-12) from *P. merguiensis*, *P. silasi*, *P. indicus*, *P. monodon*, *P. semisulcatus* and *Metapenaeus*, cleaved by *Rsa*I. (N = sample size)

Species	Fragment length (bp)	Expected fragment size from *Rsa*I	N
P. merguiensis (SREE1)	876	142, 152, 240, 360	1
P. silasi (NKE8)	900	300, 582	1
P. indicus (TA5)	876	282, 248, 334, 12 (missing band)	1
P. monodon[1]	850	295, 555	1
P. semisulcatus[2]	900	376, 524	1
Metapenaeus affinis[3]	900	100, 800	1

[1] sequence retrieved from GenBank (Acession number: AF014377)

[2] sequence retrieved from GenBank (Acession number: AF279831)

[3] sequence retrieved from GenBank (Acession number: AY264889)

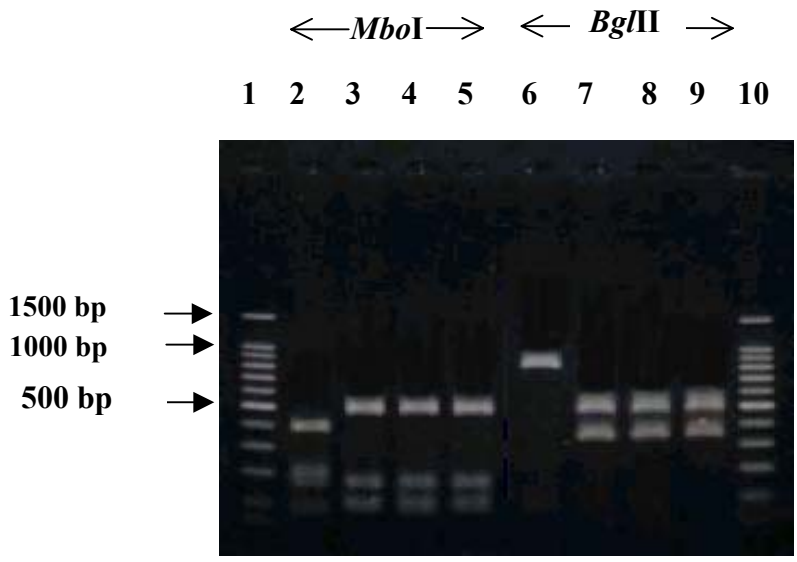

Figure 28. The restriction pattern of a 5' end fragment of COI cleaved by *Mbo*I and *Bgl*II. Lanes 1 and 10 = 100 bp markers; lanes 2 and 6 = STW4; lanes 3 and 7 = SKE26; lane 4 and 8 = TDE2; lanes 5 and 9 = SRE30.

Table 23. The restriction patterns of the 5' region of a COI fragment (800 bp) cleaved with *Bgl*II and *Mbo*I. Samples from Andaman Sea (STW, TRW and RNW), and samples from the Gulf of Thailand (SKE, SRE, NKE and TDE).

Sample	Enzyme *Bgl*II		Enzyme *Mbo*I	
	Haplotype	Fragment size (bp)	Haplotype	Fragment size(bp)
STW4	1	800	3	400, 200, 180, 20 (missing band)
STW31	1	800	3	400, 200, 180, 20 (missing band)
TRW7	1	800	3	400, 200, 180, 20 (missing band)
TRW11	1	800	3	400, 200, 180, 20 (missing band)
RNW55	1	800	3	400, 200, 180, 20 (missing band)
RNW66	1	800	3	400, 200, 180, 20 (missing band)
SREE1	1	800	3	400, 200, 180, 20 (missing band)
SREE2	1	800	3	400, 200, 180, 20 (missing band)
STWA5	2	500, 350	4	500, 180, 120
STWA6	2	500, 350	4	500, 180, 120
STWA7	2	500, 350	4	500, 180, 120
SKE26	2	500, 350	4	500, 180, 120
SKE43	2	500, 350	4	500, 180, 120
SKEB1	2	500, 350	4	500, 180, 120
SRE30	2	500, 350	4	500, 180, 120
NKE1	2	500, 350	4	500, 180, 120
NKE9	2	500, 350	4	500, 180, 120
TDE2	2	500, 350	4	500, 180, 120
TDE12	2	500, 350	4	500, 180, 120