Chapter 2: Molecular systematics of the freshwater prawn genus *Macrobrachium* Bate, 1868 (Crustacea: Decapoda: Palaemonidae) inferred from mtDNA sequences, with emphasis on East Asian species

INTRODUCTION

Freshwater prawns of the genus *Macrobrachium* Bate, 1868 (Crustacea: Palaemonidae) is a highly diverse group of decapod crustaceans, has originated from marine ancestors and subsequently migrated towards freshwaters in more than one wave. Hence, its members are known to inhabit the entire range of habitats from purely marine areas to inland hill streams and impounded waters (Tiwari 1955, Shokita 1979, Jalihal et al. 1993). To date, approximately 210 species are recognized (Short 2004); and there are numerous yet undescribed cryptic species (Chace and Bruce, 1993; Wowor and Choy 2001, Cai and Ng 2002, Cai et al. 2004, Short 2004, D. Wowor, pers. comm., Y. Cai, pers. comm.).

The genus *Macrobrachium* can be ecologically separated into two groups: most species are widely distributed and require a certain saline concentration (i.e., 10‰~35‰) to complete their larval development, as euryhaline species; others are land-locked species, with limited distributions and complete their entire life cycle in freshwater (Holthuis 1950, Johnson 1973, Shokita 1979). As *Macrobrachium* migrated towards freshwaters, the prawns gradually evolved several adaptive features; with one of them is the abbreviated larval development by reducing both the number of larval stages and the duration of the larval period (Shokita 1979, Jalihal et al. 1993). The abbreviated development of larvae in land-locked species was suggested to be a result of selective pressures for becoming established in freshwater environments (Shokita 1979, Magalhães and Walker 1988), and is a multiple convergent phenomenon overriding phylogenetic relationships even above the generic level (Magalhães and Walker 1988). On the contrary,

Pereira and Garcia (1995) suggested that since primitive palaemonids like *Troglocubanus*, *Palaemonetes* and *Pseudopalaemon*, possess abbreviated development, which could be considered a primitive trait, and the abbreviated development took place early in the origin of the family Palaemonidae, rather than being a recent process.

Because of the conservativeness of morphological characteristics, much debate has surrounded the systematic relationships of many species within this group and its taxonomy and phylogenetic inferences have until recently been based exclusively on comparisons of external morphological characters (Holthuis 1950 1952, Johnson 1973, Pereira 1997). Some species groups were proposed based on morphological similarities (synapomorphy), mainly on the rostrum and the second pereiopod (Holthuis 1950, Johnson 1973). The phylogenetic significance of these groupings remains to be tested. Apart from their taxonomy, the phylogenetic affinities among freshwater prawns are poorly understood. Pereira (1997), based on the morphological characters, carried out the first phylogenetic study on the family Palaemonidae. In recent years, Murphy and Austin (2002 2003 2004 2005) published a series of results for the phylogeny of *Macrobrachium* species based on the mtDNA fragment of the large subunit (16S) rRNA gene marker. Their studies showed some interesting results and led to the generic clarification of some local species.

The East Asia has a larger landmass within the tropics and subtropics, has a longer history of overland connection with tropical areas having rich species pools (Hamilton 1983, Guo et al. 1998). This region exhibits high species diversity of plant taxa (Guo et al. 1998, Qian and Ricklef 2000), and its freshwater fish fauna also represent one of the richest ichthyofauna in the Sino-Indian region (Koltelat 1989, Banarescu 1991). The climatic changes and geographic heterogeneity had played major roles in the diversification and speciation of East Asia's biota.

Therefore, East Asia is a superior continental model for studying increases in regional diversity through allopatric speciation (Qian and Ricklef 2000).

Previous studies of the freshwater prawns in East Asia comprised regional species surveys or zoogeographic distributions, including mainland China (Yu 1936, Dai 1984, Liang 1986, Liu et al 1990), Taiwan (Hwang and Yu 1982 1983), and Japan (Hayashi 2000a b c). There are about 37 species found in Taiwan, mainland China (Shy and Yu 1998, Li et al. 2003) and Japan (Hayashi 2000a b c); there is also a relativity high level of endemicity. Some population studies of *M. nipponense* were discussed based on allozyme variations and reproductive traits (see Mashiko and Numachi 2000 for review). Shokita (1979) inferred the speciation and origin of land-locked freshwater prawns based on prawns from the Ryukyus, and discussed the biogeography of genus *Macrobrachium* with special reference to larval dispersal (Shokita 1985). However, little work has been carried out on the phylogenetic relationships of freshwater prawns in East Asia.

For this study utilizing the mtDNA 16S rRNA marker, we attempted to investigate the phylogeny and evolution of land-locked species of the genus *Macrobrachium*, based on species from the Indo-West Pacific region, and using sequences available from GenBank. Then, we focused on species distributed in East Asia, including mainland China, Taiwan, the Ryukyus, and Japan. We used a combination of 16S rRNA and fragments of the cytochrome oxidase subunit I (COI) gene to elucidate the phylogenetic relationships of East Asian species; to test if speciation patterns of endemic species in the region resulted from multiple lineages or from a single event, and to reveal any cryptic species that are difficult to distinguish using more traditional techniques (Kowlton 2000, Hendrixon and Bond 2005, Ellis et al. 2006).

MATERIALS AND METHODS

Collection of materials

In total, 238 specimens, representing 34 putative and four undescribed species of *Macrobrachium,* were collected from East and Southeastern Asia for sequence analyses. Moreover, for 15 species, multiple individuals were sequenced from geographically distant populations to assess the monophyly of the putative species (Sites and Marshall 2003, Peters et al. 2005). One to eight specimens were analyzed per locality (Table 1). Specimens used in the present study (Table 1) were caught from the wild and preserved in 75%~95% ethanol. Five species of three closely related genera in the same family (the Palaemonidae), namely *Exopalaemon modestus, E. orientis, Palaemonetes sinensis, P. atrinubes,* and *Palaemon siuenus,* together with an atyid shrimp, *Caridina pseudodenticulata,* were included in the study as outgroup. Additional mtDNA 16S sequences, available from the GenBank, were included in this analysis (Table 2), to encompass a total of 62 species in the genus *Macrobrachium*.

DNA extraction, PCR amplification and sequencing

Total genomic DNA was extracted from abdominal muscle by proteinase K/SDS dissolution, phenol-chloroform extraction and ethanol precipitation according to standard procedures (Sambrook et al. 1989). Fragment of two mitochondrial genes, the large subunit (16S) rRNA gene and the cytochrome oxidase subunit I (COI) gene, were amplified from total gDNA by polymerase chain reaction using conserved primers: 1471 (5'-CCTGTTTANCAAAAACAT-3') and 1472 (5'-AGATAGAAACCA ACCTGG-3') (Crandall and Fitzpatrick 1996) for 16S rRNA; COI-a (5'-AGTATAAGCGTCTGGGTAGTC-3') and COI-f (5'-CCTGCAGGAGGAGGAGAC CC-3') (Palumbi and Benzie 1991) for the COI gene. The primers of 16S and COI did not work well for some species, so a new primer pair, 1471B (5'-CCTGTTTANCAAA AACATGTCTG-3') and 1472B (5'-AGATAGAAACCAACCTGGCTCAC-3'), was modified for the 16S rRNA gene,

and the new COI-fR (5'-CGTCGTGGTATGCCDTTTARWCCTA-3') primer replaced the primer COI-a.

The amplification (50 µl) for 16S rRNA contained 1 mM of each primer, 0.2 mM of each dNTP, 1 unit of *Taq* polymerase (Promega), template DNA (50~100 ng), and 1X amplification buffer containing 1.5 mM $MgCl_2$ and the primer pair of either (1471 +1472) or (1471B + 1472B). The amplification conditions involved an initial cycle with denaturation at 94 °C for 5 min, and then 35 cycles of denaturation at 94 °C for 1 min, annealing at 48~55 °C for 1 min, and extension at 72 °C for 1 min, followed by a final extension at 72 °C for 5 min. For the COI gene, the amplification used either primer pair (COI-f + COI-a) or (COI-f + COI-fR). A similar profile to that of 16S rRNA was employed except that the annealing temperature ranged from 45 to 50 °C. The size and quality of PCR products were visualized on 1.5% agarose gels.

Prior to sequencing, PCR products were purified using the gel purification kit according to the manufacturer's instructions (QIAGEN). In order to control sequence accuracy and to resolve any ambiguous bases, sequences were obtained from both directions using the same primer pairs for PCR by cycle sequencing using the ABI PRISM Dye-Terminator Sequencing Kit (Applied Biosystems) and electrophoresis on an Applied Biosystems Automated Sequencer (model 377 or 3100).

Data analyses

Forward and reverse sequences for an individual were edited using SeqMan (DNASTAR, LaserGene). All sequences were aligned using MegAlign (DNASTAR, LaserGene), with checks and adjustments made by eye using BioEdit v5.0.9 (Hall 1999). Exploratory data analysis of sequences was performed using MEGA version 2.1 (Kumar et al. 2001) and DnaSP 4.00 (Rozas et al. 2003). Pairwise sequence comparisons provided an assessment of levels of saturation by plotting the number of transitions and transversions against the uncorrected proportional

distances (p-distances) for each pair of unique sequences of *Macrobrachium* species (Morrison et al. 2004).

The 16S rRNA and COI sequences were analyzed independently and combined for species which were available for both genes. Sequences of both 16S rRNA and COI genes were combined as both genes are in effect linked and that is an appropriate way of dealing with random topological differences that are attributable to sampling error (Hipp et al. 2004). In order to test for the consistency of phylogenetic signals in the data, phylogenetic relationships were inferred using four different analytical approaches. Maximum parsimony (MP) (Camin and Sokal 1965) analysis was conducted assuming equal weightings for all characters; results were compared when gaps were treated as missing data and as a fifth character state. Neighbor-joining (NJ) (Saitou and Nei 1987), and maximum likelihood (ML) (Felsenstein 1981) analyses were estimated using an appropriate DNA substitution models calculated with ModelTest version 3.5 (Posada and Crandall 1998). In MP, the heuristic search option was used with tree-bisection-reconnection (TBR) branch swapping and 100 stepwise random additions of taxa. In NJ and MP, branch support were assessed using bootstrap resampling (Felsenstein 1985) with 1000 replicates and the full heuristic search algorithm to evaluate the reliability of the inferred topologies. Bootstrap resamplings were run with the “fast” stepwise addition algorithm and 100 replicates for ML, because of the large number of taxa involved and computational time requirement. NJ, MP and ML performed with PAUP* version 4.0b10 (Swofford 2000). Bayesian analyses (BI) were performed with MRBAYES version 3.0 (Ronquist and Huelsenbeck 2003) using the models selected by MrModeltest (Nylander 2004). The Markov chain Monte Carlo (MCMC) chains were run for one million generations, and trees were saved each 100 generation (the first 1000 trees were discarded as “burnin”). In BI, posterior probabilities are true probabilities of clades and those with values of 95% or greater deemed to be significantly supported.

RESULTS

Sequence characteristics and variations

The 16S rRNA sequence amplified by (1471B + 1472B) varied from 524 to 533 bp in length. Additional sequences obtained from GenBank were shorter, final truncated lengths for the multiple alignments were 442 bp, including 20 sites with gaps, 191 sites were variable, with 160 were parsimony informative. Numbers of transitions outnumbered transversions in all comparisons by a factor of approximately 1.7. The COI sequence amplified by (COI-f + COI-fR), contained 608 bp, in which 276 sites were variable, and 258 were parsimony informative. No stop codon was revealed when the COI sequences were translated into amino acids. The numbers of transitions outnumbered transversions with an average of 2.2. Base frequencies in both mtDNA genes showed an AT bias (G+C content: 35.8% and 40.7% for 16S and COI, respectively). The combined data set of 16S+COI multiple alignments was 1142 bp in length including 20 sites with gaps, 472 variable sites, and 410 parsimony informative sites. All mtDNA sequences determined in this study were deposited in GenBank/DDBJ databases under accession numbers DQ194904~DQ194973 and AB235240~AB235308 (Table 1).

In the gene-saturation analyses, for 16S, substitutions of transitions and transversions were approximately linear in distribution with a positive slope for the regression (R^2 = 0.714 and 0.660 for transitions and transversions, respectively, Fig. 1), indicating that 16S rRNA is not saturated. For the COI data, transitions and transversions were plotted by separate codon position, a saturation tendency was shown for transitions in the 3rd position, not for the 1st and 2nd positions, which appeared to reach a plateau at p-distances above 20% (Fig. 2). Sequence divergence estimates among the *Macrobrachium* species ranged 0.47%~22.44% and 0.16%~25.54% for 16S and COI, respectively; for conspecific individuals from the same locality, they ranged 0.00%~0.11% and 0.00%~3.70% for 16S and COI, respectively. Between

populations (or individuals) from different localities of conspecifics, they ranged 0.00%~3.20% and 0.00%~12.63%, (the most-distant populations of *M. grandimanus* from the Ryukyus and Hawaii in this study) for 16S and COI, respectively. Four undescribed species were confirmed to be genetically distinct from other species, with interspecific divergence (3.5%~19.9% and 9.48%~27.67% for 16S and COI, respectively). These will hereafter, be referred to as *Macrobrachium* sp. 1, sp. 2, sp. 3, and sp. 4. We also found some discordant cases in which intraspecific divergences (4.9%~9.2% and 13.23%~17.24 for 16S and COI, respectively), greatly exceeded the intraspecific divergences ranges, and might therefore reflect interspecific differences. Such discordant situations were found in *M. latidactylus*, *M. latimanus*, *M. jaroense*, *M. placidum*, and *M. equidens* for samples taken from multiple localities. According to these results, we believe such specimens may represent "cryptic species" (sensu. Gusmão et al. 2000, Knowlton 2000, Hendrixon and Bond 2005, Ellis et al. 2006). These species were thus referred to as *M.* sp. 5, sp. 6, sp. 7, and sp. 8 (Figs. 3, 4 and 5), the species *M. equidens* is for further discussion. In contrast, the divergence between two morphologically distinct species, *M. formosense* and *M. hainanense* (0.01%~0.02% and 0.16%~0.66% for 16S and COI, respectively), was as close as that of the intraspecific level. This was also found in the species *M.* cf. *horstii* and *M. placidulum* (0.01% and 0.66% for 16S and COI, respectively).

Phylogenetic analyses

Based on results from Modeltest, the best-fit models in the NJ and ML were as follows: 16S, HKY+I+G (Hasegawa et al. 1985), with correction for among-site rate variation estimation (G) = 0.6420 and proportion of invariable sites (I) = 0.4499. COI: GTR+I+G (General Time Reversible model; Rodriguez et al. 1990), with correction for G = 0.4023 and I = 0.4704. For the combined dataset: TrN+I+G (Tamura-Nei model; Tamura and Nei 1993) with correction for G = 0.7019 and I = 0.5354. For BI the best-fit models selected by Akaike Information Criterion (AIC) in

MrModeltest were as follows: 16S, GTR+I+G, with correction for G = 0.5897 and I = 0.4106. COI: GTR+I+G, with correction for G = 0.4133 and I = 0.4770. For the combined dataset: GTR+I+G with correction for G = 0.6176 and I = 0.4675.

For 16S, 103 haplotypes for *Macrobrachium* species, including four undescribed species and seven species of outgroup were analyzed. A 50% majority consensus tree was obtained from the MP analyses. All the four phylogenetic analyses generated similar tree topologies when gaps were treated either as a fifth character or as missing data (Fig. 3). The major differences between the four analyses lay in the levels of support provided for the various clades. Tree topologies support the monophyly of the genus *Macrobrachium* with high bootstrap values in NJ, MP and BI. The outgroup genera *Palaemonetes*, *Palaemon*, and *Exopalaemon* formed a monophyletic clade with high bootstrap support, and were paraphyletic to *Macrobrachium*. *Exopalaemon modestus* and *E. orientis* formed a sister taxon pair with high bootstrap support in all analyses. *Palaemonetes atrinubes* and *P. sinensis* formed a sister taxon pair with weak bootstrap support, *Palaemon serenus* and *M. intermedium* formed a sister taxon pair with bootstrap support. The species *M. intermedium* was located outside the *Macrobrachium* clade, being a species more-closely related to *Palaemon serenus* than to any other species.

The deeper internal nodes were generally unresolved in these analyses. Some species groups, with good support for many terminal clades, were revealed giving some support to earlier classifications except the *M. equidens* species group. In addition to the Central and South American and West African species clades as reported by Murphy and Austin (2005), another four clades were revealed. *Macrobrachium rosenbergii* formed a species group together with two Indian euryhaline species, *M. gangeticum* and *M. malcolmsonii*, and two Indian land-locked species, *M. lamarrei* and *M. sankollii*. The next two distinct groups were endemic to East Asia; one was the land-locked *M. asperulum* species group, containing *M. asperulum, M. anhuiense, M. pinguis, M. shokitai*, and *M. maculatum*. *M. anhuiense* and *M. pinguis*, formed sister species with

populations of *M. asperulum* distributed in North and South Taiwan, respectively. Another species group containing the euryhaline species *M. formosense* and *M. hainanense* and the land-locked species, *M. nipponense, M. inflatum*, and an undescribed species, *M*. sp. 4 was placed at the basal position and is distributed in southeastern mainland China. The fourth group includes several widely distributed euryhaline species, namely *M.* cf. *horstii*, *M. placidulum*, *M. placidum*, and one cryptic species, *M*. sp. 8, which is morphologically very close to *M. placidum*.

Multiple samples from distant geographic populations of putative species were grouped into species-specific monophyletic groups with low to high bootstrap support (bootstrap values 51~100), with the exception of two species. The *M. equidens* species group, suggested by Johnson (1973), including *M. equidens,* M. *idae, M. mammillodactylus,* and *M. novaehollandiae*, did not form a monophyletic group. Moreover, specimens of *M. equidens* from four different localities, those of Taiwan, the Philippines, and Australia formed a lineage and were separated, with distinct genetic distances (16.7%~17.4%), from specimens collected from Singapore, the type locality of *M. equidens*. Such disagreement was more evident in other species. Two populations (I, II) of *M. hainanense* from mainland China (Sample locality Guangdong and Hainan, Table 1) have an inter-population distance of 0.08%. When we compared our *M. hainanense (M. hainanense* (ML)) sequences with *M. hainanense* (HK) (accession no.: AY377841), collected from Hong Kong (Murphy and Austin 2005), there were inconsistencies with a significant distance of 7.6%. The *M. hainanense* (HK) sequence was placed into a different clade with land-locked species and was closely correlated to *M. maculatum* with a divergence of 3.8%, wherein our *M. hainanense* sequences formed a clade with the euryhaline species group and was closely correlated with *M. formosense*.

Species containing only 16S rRNA sequences in the previous studies or only available in GenBank (Table 2) were not included in constructing COI phylogenetic trees (Fig. 4). Most of the species groups revealed above, including the four undescribed ones, four "cryptic species",

and the incongruence of *M. equidens* and *M. hainanense*, were supported with similar topologies as constructed in the phylogenetic analyses of the 16S sequences. The monophyly of the *Macrobrachium* species clade cannot be confirmed by the COI data, as outgroup species, *Exopalaemon modestus* and *E. orientis*, nested as one of the polyphyly clades of *Macrobrachium* species taxa. The non-monophyletic structure of the *Macrobrachium* species may be attributed to the saturation of 3rd codon of COI gene (Fig. 2).

Analysis of the combined dataset

Although the deeper internal nodes were generally unresolved and the relationships among the species were not well-resolved by the combined sequences, certain phylogenetic relationships of the species complexes however were shown to be well supported (Fig. 5). For example, the morphologically similar land-locked species endemic to East Asia, including the *M. asperulum* species group *(M. anhuiense*, *M. pinguis*, and *M. shokitai*) and other land-locked species of *M. edentatum* and *M. maculatum* and two undescribed species *M*. sp. 2, *M*. sp. 3 formed a monophyletic group. The morphologically dissimilar land-locked species *M. fukiense* and Southeast Asian species, *M. malayanum*, *M. platycheles*, and *M. yui*, were not included. *M. anhuiense* and *M. pinguis* formed sister pairs to *M. asperulum* distributed in northern and southern Taiwan, respectively, with high bootstrap support.

The species group containing *M.* cf. *horstii*, *M. placidulum*, and *M. placidum* did not form a monophyletic clade with the morphologically similar *M. lepidactyloides* (Holthuis 1950) in either 16S or COI analysis separately, but did form a monophyletic group when the combined dataset was analyzed. Another species group with similar morphology of unequal second periopod, containing *M. esculentum*, *M. lanatum*, *M. latidactylus*, and *M.* sp. 5 formed a clade. One cryptic species, *M.* sp. 8, was nested in this group. The incongruence of non-monophyly of *M. equidens* and *M. hainanense* were also shown.

DISCUSSION

Phylogenetic relationships

Although we have included more species distributed in East Asia and the Indo-West Pacific region than Murphy and Austin (2005), the phylogeny of *Macrobrachium* species based on mtDNA 16S rRNA still showed a poor resolution of 'starburst' relationships, which lack internal structure with short internal branch lengths and longer tips among species of *Macrobrachium* (Fig. 3). Such phylogenetic relationships may have been caused by a weakness of the marker when it reaches saturation or by a lack of power by the data to resolve relationships among taxa (Albertson et al., 1999). 16S rRNA, used in this study, is not saturated (Fig. 1). Alternatively, the unresolved phylogeny detected herein may also be explained by a rapid radiation, as suggested in the early studies of *Sebastes* rockfishes (Johns and Avise 1998), fairy shrimp (Daniels et al. 2004), Caribbean sponge-dwelling snapping shrimp (Morrison et al. 2004), squat lobsters (Machordoma and Macpherson 2004) and freshwater crayfish (Shull et al. 2005). In all of those cases, resolution and/or support for the nodes in question were poor, suggesting a real phenomenon resulting from rapid radiation, rather than a simple paucity of appropriate data. When we focused on the prawn species distributed in East Asia, the COI sequences revealed the same pattern of phylogenetic relationships of poor resolution at the internal nodes (Fig. 4). The combined dataset of 16S and COI sequences allowed the current study to show a relatively clear and well-supported pattern of phylogenetic relationships among some species groups (Fig. 5).

Although we were unable to reconstruct detailed relationships of all *Macrobrachium* species due to their rapid radiation, whereas relationships among many of the terminal taxa were moderately to well support which have revealed several important features. The monophyly of the genus *Macrobrachium* with outgroup was supported by phylogenetic analyses (Fig. 3, 5), *M. intermedium* was excluded from the genus *Macrobrachium* at earlier studies (Pereira 1997,

Murphy and Austin 2002 2003, Short 2004). Multiple origins of *Macrobrachium* fauna on various continents (or regions), like Central and South America, East Asia, the Indo-West Pacific, and India, are suggested, supporting the results of Murphy and Austin (2005). The planktonic larval stage with salinity tolerance may play an important role in long-distance dispersal and may have contributed to the widespread distribution of species (Shokita 1985, Jalihal et al. 1993, de Bruyn et al. 2004), such as *M. latidactylus*, *M. latimanus*, *M. grandimanus*, *M. rosenbergii*, and *M. lar*. Minor genetic differences among conspecific, widespread euryhaline populations over broad geographic areas suggest that gene flow has been continuously ongoing, and that they tend to have low levels of differentiation compared to species with non-planktonic larval phase (Cameron 1986). Not all of the land-locked freshwater prawn species, including *Exopalaemon modestus* and *Palaemonetes sinensis*, formed a monophyletic group, or was located in the basal position, suggesting that they did not diverge from a single marine ancestor, but likely have originated from marine ancestors and subsequently migrated towards freshwater in multiple waves of migration (Tiwari 1955, Jalihal et al. 1993). An abbreviated larval development pattern (ovigerous female with large eggs) has been suggested as being a process resulting from selective pressures caused by attempts to become established in freshwater environments, and is a result of adaptive convergence (Shokita 1979, Magalhães and Walker 1988). This pattern is parallel to another freshwater shrimp in the family Atyidae (Magalhães and Walker 1988) and is also in Jamaican's *Sesarma* crabs (Schubart et al. 1998). Jayachandran's (2001) suggestion that the genus *Macrobrachium* could be grouped into two categories based on morphological characters of the second pereiopod is not supported by our findings. Our results (Fig. 3, 4 and 5) demonstrate that the most-common morphological characters used in the taxonomy of the genus *Macrobrachium* (e.g., the second pereiopod or the rostrum), do not form into a monophyletic group, and do not always have phylogenetic value. The morphological characters (such as the unequal second pereiopod, big robust claws, spine, etc.) should have developed independently

through the invasion of inland waters.

Taxonomic implications

Based on mtDNA sequences, 11 of 15 species obtained from geographically distant populations formed monophyletic lineages; four undescribed species (*M.* sp.1~sp. 4) were identified, and five cryptic species (*M.* sp. 5~sp. 8 and *M. equidens*), grouped with *M. latidactylus, M. latimanus, M. jaroense, M. placidulum* and *M. equidens,* were inferred according to the phylogenetic reconstructions (Figs. 3, 4 and 5) and sequence divergence levels.

Some misidentification or invalid species were revealed. The species *M. hainanense* (accession no.: AY377841), collected from Hong Kong by Murphy and Austin (2005) was inconsistencies to our specimens from two China's populations (Sample locality Guangdong and Hainan Provinces, Table 1) with a significant genetic distance and different position in phylogenetic tree (Fig. 3). Holthuis (1950) commented that *M. hainanense* was so closely related to *M. formosense*, that it perhaps should only be considered as a subspecies. According to the divergence and phylogenetic tree (Fig. 3), *M. hainanense* used by Murphy and Austin (2005) is most probably a "misidentified species". Examination of Hong Kong specimens shows that there is an undescribed species of the *Macrobrachium asperurum* group in Hong Kong. The subadult specimens of this species could be easily confused with that of *Macrobrachium hainanense*, which also occurs in Hong Kong. *M. anhuiense* and *M. pinguis* formed sister pairs to *M. asperulum* distributed in northern and southern Taiwan, respectively, with high bootstrap support and showed intraspecific level genetic divergences in all of the 16S, COI and combined dataset analyses. It has been suggested that *M. pinguis* is an invalid species described on the basis of undeveloped males, and was synonymized with *M. asperulum* by Liu et al. (1990). *Macrobrachium anhuiense* was also excluded from the species list of Chinese palaemonoid fauna (Li et al. 2003). The present DNA data supports both taxonomic actions. Both species,

based on both morphological, phylogenetic evidences, could not be separated from *M. asperulum*, and should be treated as invalid species.

When the combined dataset was analyzed, species group, containing *M.* cf. *horstii*, *M. placidulum*, and *M. placidum*, forms a monophyletic group. Molecular studies did not support the close relationship of *M. lepidactyloides* and *M. placidum* revealed by morphological similarity. This contradicts suggestion by Chace and Bruce (1993) that *M. lepidactyloides* may be synonymous with *M. placidum* in morphology. The species *M.* cf. *horstii* (cf Shy and Yu 1998) is closely allied to *M. placidulum* with intraspecific level of genetic divergence and it is also morphologically similar to the latter species, should be regarded as conspecific with *M. placidulum*.

The 5 cryptic (*M.* sp. 5, sp. 6, sp. 7, sp. 8 and the *Macrobrachium equidens*) species all showed minor morphological differences of "intraspecific" variation, but were genetically very distinct from other populations (localities), with value of interspecific divergences and closely allied with high bootstrap values. This also suggests that the use of traditional morphological characters alone cannot be used to accurately diagnose natural species groups of *Macrobrachium* (Holthuis 1950, 1952, Johnson 1973).

In this study, the intraspecific 16S sequence divergence estimates between populations (or individuals) from different localities ranged 0.0%~3.2%. The significant divergence of 16S (5.1~6.2%) between eastern and western *M. rosenbergii* clades along Huxley's line (de Bruyn et al. 2004) is far beyond the ranges of intraspecific divergence when compared with *M. lar*, *M. latidactylus*, *M. mammillodactylus*, *M. latimanus*, and *M. grandimanus* which are also distributed across Huxley's line, in present study. Genetic evidences of present study supports the previous studies (De Man 1879, Johnson 1973, Lindenfelser 1984, Malecha 1987, Wowor and Ng 2001)

that *M. rosenbergii* may actually represent two distinct species, i.e. one eastern and the other western forms.

Macrobrachium is a notoriously difficult genus taxonomically, as the morphological plasticity of taxonomically important traits (e.g., the rostrum and/or the second pereopod) change so much and so gradually during their growth (Holthius 1950) and influenced by environmental parameters (Dimmock et al. 2004). Morphologically similar species are often quite distinct genetically. However, this might not be reflected in the phylogenetic relationships among them, as shown in the *M. equidens* species group (Johnson 1973), suggesting that conservative systematic traditions or morphological stasis may be involved (Knowlton 2000). Most genetic analyses of species boundaries, in marine crustaceans (see Knowlton 2000 for review) and freshwater macroinvertebrates (Baker et al. 2004, Shih et al. 2004 2005 2006), confirm or reveal the existence of cryptic species that are difficult to distinguish using more traditional techniques. Some cryptic species are distinguished by surprisingly large genetic differences (Kowlton 2000, Hendrixon and Bond 2005, Ellis et al. 2006). This problem highlights a number of features of the species group in these analyses. First, despite the diversity of species of the genus, they are all relatively conservative in general appearance, and taxonomical mistakes are easily made. Second, it shows the considerable value of having multiple samplings within each taxonomic group of interest, so that possible errors can be detected (Sites and Marshall 2003, Peters et al. 2005). It also suggests that using a single example of an individual or population to represent

species in the overall analysis should be treated with caution and is not justified. It is necessary to reevaluate the practical species concept based on such a multiple-sample analysis. The cryptic species detected here suggest that molecular techniques will be a significant help in delimiting species and understand their relationships (Knowlton 2000, Hendrixon and Bond 2005, Ellis et al. 2006).

Some of the species groups, including species through different geographic regions, could be suggested from Central and South America, India, and East Asia. *Macrobrachium rosenbergii*, the widely distributed euryhaline species (De Man 1879, Johnson 1973) with an extended type of larval development and regarded as probably fairly "ancient" in nature by Johnson (1973), was not placed at the basal position of the *Macrobrachium* species clade. It formed a species group, suggested by Johnson (1973), together with two Indian euryhaline species, *M. gangeticum* and *M. malcolmsonii*, and two land-locked species endemic to India, *M. lamarrei* and *M. sankollii*. Species groups containing such a mixture of species with different life cycles may suggest that the species group may have evolved from a single ancestral lineage. Such a lineage was also found in the *M. nipponense*/ *M. formosense*/ *M. hainanense* species group, and the group with the asymmetrical second pereiopod, i.e., *M. lanatum*, *M. esculentum*, *M. latidactylus*, and cryptic species *M.* sp. 5. The species *M. handschini*, a land-locked species in Australia, formed a sister pair with *M. esculentum* (specimens collected from Taiwan and Philippines).

An endemic speciation event in East Asia was suggested in the land-locked species group, including the *M. asperulum* species group (Fig. 5). This land-locked species group did not form a monophyletic group with the land-locked species endemic to Southeast Asia (*M. yui*, *M. malayanum*, and *M. platycheles*), which implies its single lineage. Among them, *M. asperulum* is the most widely distributed species, being known from south Siberia to southeastern China

(Holthuis 1950). The other land-locked species have restricted distribution ranges. For freshwater-dependent prawns, factors responsible for dispersal generally involve land continuity and therefore river confluences (e.g., during sea-level lowering in glacial maxima), as well as river capture in headwaters (Banarescu 1991). The best explanation is that it represents fragmentation of a widespread ancestral taxon (vicariance) through allopatric speciation (Qian and Ricklef 2000), rather than a dispersal phenomenon from a more-restricted "center of origin" (Wiley 1988).

In addition to the ancient radiation speciation, there is evidence for ongoing freshwater invasion and recent speciation in East Asia. The processes of freshwater invasion and penetration of cool temperate areas (at high latitudes) may be represented by *M. nipponense*. *Macrobrachium nipponense* is distributed along the coastline from northern Southeast Asia north to East Asia and Japan (Liu et al. 1990, Cai and Dai, 1999). Its narrow salinity tolerance, and relatively shorter zoeal period may have limited the dispersal distance to a dispersal pattern of from estuary to estuary, apparently not a trans-oceanic dispersal of which larval stages of euryhaline species could dispersal across the ocean by ocean currents as *M. gradimanus* (Shokita 1985) and *M. rosenbergii* (de Bruyn et al. 2004). Some of *M. nipponense* populations, on the basis of allozyme variation and reproductive traits, split into freshwater and estuarine populations in the same river in Japan (see Mashiko and Numachi 2000 for review), while the populations in mainland China and Taiwan are now land-locked species (Liu et al. 1990, Cai and Dai, 1999). This probably represents different steps in the process of inland water invasion. The land-locked populations of *M. nipponense*, in mainland China and Taiwan, represent an advanced state of freshwater invasion, while populations of *M. nipponense* in Japan may represent both the advanced state (the land-locked populations inhabiting freshwaters) and euryhaline. *M. hainanense* and *M. formosense* are closely related species in morphology; Holthuis (1950) commented that that *M. hainanense* perhaps should only be considered a

subspecies. These two species exhibit the lowest interspecific DNA divergence, and closely relationship (Fig. 3, 4 and 5). Based on the fact that *M. formosense* has restricted distributions than *M. hainanense* (Li et al. 2003); we believe that *M. formosense* may represent a newly derived species in recent geological times.

However, this study is only based on two mitochondrial markers. Further studies, particularly using other mitochondrial or nuclear sequence data and including more species, are obviously required to further investigate their evolutionary history. New data on the biology and ecology of these species and their habitats, and updated knowledge of the paleogeographical history will help to clarify the possibility of microhabitat and behavioral specialization giving rise to the radiation of *Macrobrachium*.