Revised Phylogeny of Extant Xiphosurans (Horseshoe Crabs)

B. Akbar John, Hassan I. Sheikh, K.C.A. Jalal, K. Zaleha and B.Y. Kamaruzzaman

Abstract An attempt was made to revise the molecular phylogeny of extant xiphosurans (Horseshoe crabs) using universal barcode gene cytochrome oxidase C subunit 1. All four extant horseshoe crab species namely *Limulus polyphemus* (American horseshoe crab), *Tachypleus gigas, T. tridentatus* and *Carcinoscorpius rotundicauda* (Asian conspecifics) together with predicted ancestral lineages (insects, scorpions and common crabs) were considered for phylogram construction using distance matrix methods. Genetic distance (GD) data analysis revealed the distant genetic relatedness of *L. polyphemus* with Asian conspecifics. More interestingly, the monophyletic origin of *Tachypleus gigas* and *Tachypleus tridentatus* was quite evident in the phylogram which other molecular markers failed to address. Close genetic relatedness of horseshoe crabs with insects showed that they might have evolved from ancient aquatic insects. The efficiency of cytochrome oxydase C subunit 1 gene in species level identification among the horseshoe crab genome was clear in both the phylogram together with the precise identification of the differential developmental stages to the species level.

Keywords Horseshoe crabs · Living fossil · Xiphosuran · Genetic lineage · Malaysia

B. Akbar John (✉) · K. Zaleha
INOCEM Research Station (IRS), Kulliyyah of Science, International Islamic University Malaysia, Kuantan, Pahang, Malaysia
e-mail: akbarjohn50@gmail.com

H.I. Sheikh
Department of Biotechnology, Kulliyyah of Science, International Islamic University Malaysia, Kuantan, Malaysia

K.C.A. Jalal · B.Y. Kamaruzzaman
Department of Marine Science, Kulliyyah of Science, International Islamic University Malaysia, Kuantan, Malaysia

S. Trivedi et al. (eds.), *DNA Barcoding in Marine Perspectives*,
DOI 10.1007/978-3-319-41840-7_7

1 Introduction

Among the bewildering array of animal taxa, horseshoe crabs are unique in their genetic makeup and are commonly known as "Living Fossil". An Intriguing characteristic of horseshoe crabs is that they are morphologically similar in look and having virtually unchanged genetic makeup which helped them in withstanding various environmental stresses for the past 150 million years (Rudkin and Young 2009; Kamaruzzaman et al. 2011). It is interesting to note that, within the impressive diversity and ecological range of extant chelicerate arthropods, only the xiphosurid horseshoe crabs retain a primitively obligate aquatic habit harking back to their distant genealogical roots. It was noted that the fossil record of the basic xiphosurid horseshoe crab body plan has been extended back to the late Ordovician Period, about 445 million years ago, demonstrating an origin that lies outside of the paraphyletic 'synziphosurines' (Obst et al. 2012).

Global distribution pattern of four extant species of horseshoe crabs showed restricted inhabitation of *Limulus polyphemus* along the American coastline especially in Gulf of Mexico, while other three Asian conspecifics such as *Tachypleus tridentatus, T. gigas*, *Carcinoscorpius rotundicauda* are inhabiting Indo-china coastal waters from Bay of Bengal up to South Philippines. In Malaysia, their distribution was noted in both east and west coast of Peninsular Malaysia with the restricted distribution of *T. tridentatus* in Borneo Island (Sabah and Sarawak) (Akbar John et al. 2011, 2012). During spawning time adult horseshoe crabs migrate from the offshore continental shelf to spawn on intertidal sandy beaches (in case of *T. gigas*) and sandy mud beaches and mangrove area (in case of *C. rotundicauda*) during full and new moon days (Zaleha et al. 2010).

Detailed morphological descriptions of the four species of horseshoe crabs were given by Mikkelsen (1988) and the comparisons of morphological differences between species were presented by Chiu and Morton (2003). Morphological characters, including the shape of the prosoma and telson, the shapes and numbers of claspers in the male and the shapes of marginal spines in the female, have been considered important in the classification of horseshoe crabs (Sekiguchi and Nakamura 1979; Mikkelsen 1988). Chiu and Morton (2003) had extensively discussed morphological variations between *Tachypleus tridentatus* and *Carcinoscorpius rotundicauda* using conventional morphometric approach. Yamasaki et al. (1988) had studied geographic variations in body sizes (maximum prosomal width) of the four extant horseshoe crab species and he proved significant variation in prosomal width of *T. tridentatus* and C. *rotundicauda* collected from different countries. These little morphological differentiations among horseshoe crab lineages have resulted in substantial controversy concerning the phylogenetic relationship among the extant species of horseshoe crabs, especially among the three species in the Indo-Pacific region. Earlier studies suggested that the three species constitute a phylogenetically irresolvable trichotomy (Xia 2000).

These discrepancies have attracted various researchers to concentrate on their genomic structure to differentiate the species at the gene level. Hence, the present

study was aimed to revise the existing molecular phylogeny of horseshoe crabs and to predict their possible ancestry using universal barcode gene (mitochondrial Cytochrome Oxidase C subunit 1 gene) as a benchmark reference.

2 Materials and Methods

2.1 Sample Collection, Preparation and Larval Rearing

Matured horseshoe crabs (*Tachypleus gigas*) were collected from Balok and Pekan nesting grounds and a female *T. gigas* sample was collected from Pulau Gaya, Sabah, Eastern Malaysia. Mangrove horseshoe crab (*Carcinoscorpius rotundicauda*) samples were collected from Sitieu mangrove forest (Terengganu, East coast of Malaysia) during May and August 2010. All the samples were identified to the species level using conventional taxonomic keys (Yamasaki et al. 1988; Chiu and Morton 2003) Fertilized eggs of *T. gigas* were sampled from Pekan and immediately transported to Institute of Aquatrop (University Malaysia Terengganu) in an aerated condition. Eggs were kept in filtered sea water (salinity 33 ± 2 ppt) under aerated condition in a larval rearing tank for 30 days. After a month, Pretrilobite stage of *T. gigas* swimming in the amniotic fluid were sampled and preserved in 95 % ethanol for DNA sequencing. In next 2 weeks, free swimming trilobite stage was sampled from larval rearing tank and preserved in 95 % ethanol as mentioned above. Precautions were taken to avoid fungal attack on developing eggs by constantly changing the filtered sea water in every 3 days. Simultaneously matured female crabs were dissected out using sterilized scissors and forceps to collect Apodeme tissue and Immatured eggs. Samples such as immatured egg, matured egg, flesh (Apodeme), pretrilobite larvae and trilobite larvae were collected in 1.5 ml Eppendorf tube containing 95 % ethanol for DNA isolation. Prior to this, all the samples were photographed for future reference. The sample details and coordinates of sampling locations are given in Table 1.

2.2 DNA Extraction, PCR and Sequencing

Salting out procedure was adopted to extract the DNA from the samples (Ajmal Khan et al. 2010; Akbar John et al. 2010; Prasanna Kumar et al. 2011). Approximately 570 bp of COI gene from mitochondrial DNA was amplified using Forward (Fish F2_t1: TGTAAAACGACGGCCAGTCGACTAATCATAAAGA TATCGGCAC) and Reverse primer (FishR2_t1: CAGGAAACAGCTATGAC ACTTCAGGGTGACCGAAGAATCAGAA) under the PCR condition of an initial step of 2 min at 95 °C followed by 35 cycles of 0.5 min at 94 °C, 0.5 min at 54 °C and 1 min at 72 °C, followed in turn by 10 min at 72 °C and then held at 4 °C.

Table 1 Sample details and coordinates of each sampling locations

Scientific name	Type of sample	Sex	Sampling area	Latitude and longitude	
Carcinoscorpius rotundicauda	Immatured egg	F	Sitieu forest	3°36.157′ N	103°23.952′ E
Carcinoscorpius rotundicauda	Flesh	F	Sitieu forest	3°36.157′ N	103°23.952′ E
Tachypleus gigas	Flesh	M	Balok	3°56.218′ N	103°22.623′ E
Tachypleus gigas	Flesh	F	Balok	3°56.201′ N	103°22.615′ E
Tachypleus gigas	Flesh	F	Pulau Gaya	6.016666° N	116.0333333° E
Tachypleus gigas	Fertilized egg	F	Pekan	3°36.157′ N	103°23.952′ E
Tachypleus gigas	Immatured egg	F	Pekan	3°36.157′ N	103°23.952′ E
Tachypleus gigas	Pretrilobite	–	Pekan	3°36.157′ N	103°23.952′ E
Tachypleus gigas	Trilobite	–	Pekan	3°36.157′ N	103°23.952′ E
Tachypleus gigas	Immatured egg	F	Pulau Gaya	6.0166667° N	116.0333333° E

PCR products were visualized on 2 % China Agarose gel and the photographed using Gel imager Under UV light. Products were labeled using Qiagen sequencing kit and sequenced unidirectionally using a MegaBace capillary sequencer at Bioserve biotechnologies pvt. Ltd. Hyderabad, India. Generated sequences were edited using Chromas Pro 2.33v. Sequences were deposited in NCBI under the Genbank ID JF896105-JF896114 (Table 2).

2.3 *Software Analysis*

DNA sequences generated in this study together with the horseshoe crab and predicted sister taxa sequences (retrieved from public DNA data banks were run in Clustal X 2.0.6v for multiple sequence alignment under default setting (Larkin et al. 2007). Nucleotide composition was determined using BioEdit 7.0.9v (Hall 1999). Molecular Evolutionary Genetics Analysis (MEGA) beta 4.1v was used to generate phylogram using distance matrix methods such as Neighbor Joining (NJ) method and Un-weighted Pair Group *Method* with Arithmetic Mean (UPGMA) (Tamura et al. 2007). Kimura 2 Parameter (K2P) was used as a distance model to generate phylogenetic tree in both the methods (Kimura 1980). Genetic Distance (GD) data were also retrieved at each codon position using the same K2P distance model.

Table 2 Details of sequences and corresponding developmental stages submitted in National Centre for Biotechnological Information (NCBI)

Scientific name	Type of sample	Sex	Genbank accession no	Protein ID	Sequence length in bp
Carcinoscorpius rotundicauda	Immatured egg	F	JF896105	AEG75796	635
Carcinoscorpius rotundicauda	Flesh	F	JF896106	AEG75797	647
Tachypleus gigas	Flesh	M	JF896107	AEG75798	565
Tachypleus gigas	Flesh	F	JF896108	AEG75799	486
Tachypleus gigas	Flesh	F	JF896109	AEG75800	537
Tachypleus gigas	Fertilized egg	F	JF896110	AEG75801	534
Tachypleus gigas	Immatured egg	F	JF896111	AEG75802	506
Tachypleus gigas	Pretrilobite	–	JF896112	AEG75803	537
Tachypleus gigas	Trilobite	–	JF896113	AEG75804	538
Tachypleus gigas	Immatured egg	F	JF896114	AEG75805	691

Note Sequences available online from June 5th 2011 on NCBI portal

3 Results

3.1 Neighbor Joining (NJ Method)

Phylogenetic tree was constructed using Neighbor Joining method to verify the efficiency of *cox*1 gene in delineating closely related and morphologically cryptic species of horseshoe crabs (Fig. 1). Phylogram was constructed with 36 sequences (horseshoe crabs = 21; Insects = 5; Scorpion = 8 and out groups = 2). The out groups used were *Portunus pelagicus* (Class: Malacostraca) and *Artemia franciscana* (Class: Branchiopoda) distinctly segregated in separate branch in phylogram proving its reliability besides the higher bootstrap values in the internal branch nods. Four distinct clads were found in the phylogram having horseshoe crabs in Clad 1, Scorpions in Clad 2, Insects in Clad 3 and out group organisms in Clad 4. The phylogeographical signals were apparent in Clad 1, segregating *T. gigas* collected from Pulau Gaya (Borneo) from the very species sampled in East coast of Malaysia (Balok and Pekan). The efficiency of cox1 gene in species level identification of various developmental stage of *T. gigas* was noted but it was inefficient in sex determination and precise identification of various developmental stages. The constructed phylogram proved the monophyletic nature of *T. gigas* with *T. tridentatus*. Atlantic horseshoe crab (*L. polyphemus*) was genetically distinct from other species of horseshoe crabs. Insects were clumped together in Clad 3 which shared a common branch node with horseshoe crabs proving their closer genetic relatedness with horseshoe crabs. Terrestrial scorpion species used in this analysis

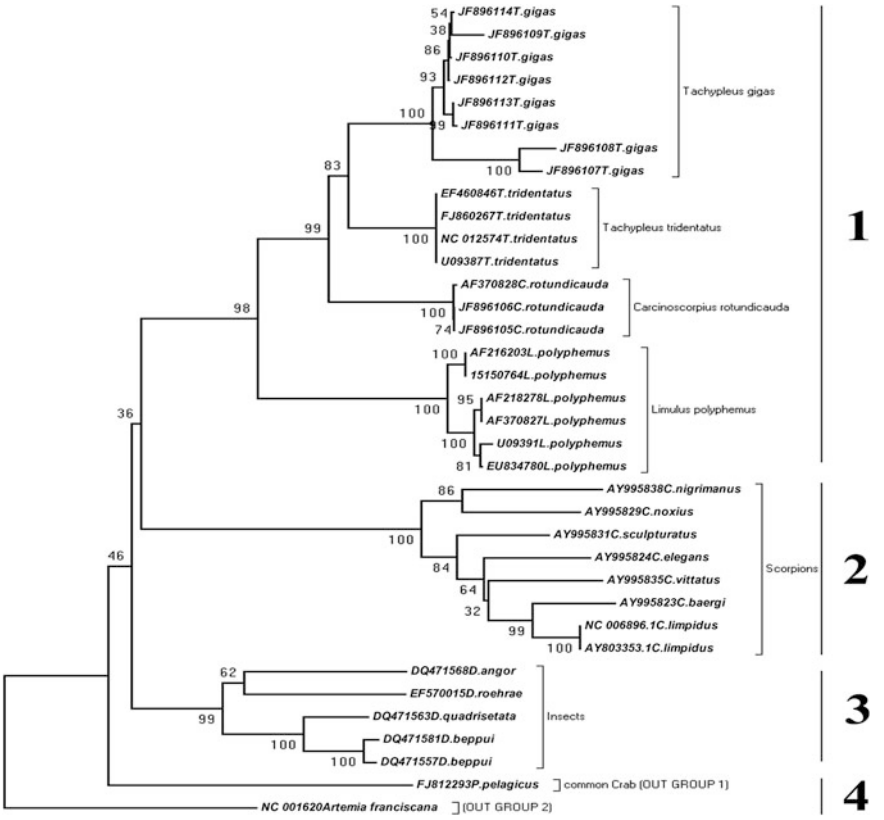

Fig. 1 Neighbor-Joining (NJ) phylogenetic tree was constructed to determine the evolutionary history of horseshoe crabs with related sister taxa (Saitou and Nei 1987)

were clumped together in Clad 2 indicating their distant genetic relatedness with horseshoe crab species. Among the horseshoe crabs, *L. polyphemus* had comparatively higher GC content than the other species of horseshoe crabs. Average GC content in *L. polyphemus* was 38.02 % followed by *C. rotundicauda* 36.49 %, *T. gigas* (33.55 %) and *T. tridentatus* (32.78 %).

3.2 *Unweighted Pair Group Method with Arithmetic Mean (UPGMA) Method*

UPGMA method was adopted to construct hierarchical clustering Phylogram to infer genetic relatedness of selected group of species (Fig. 2). This method strictly follows the molecular clock hypothesis (constant rate of evolution) and hence useful in verifying the information obtained through NJ-method. Same sequence

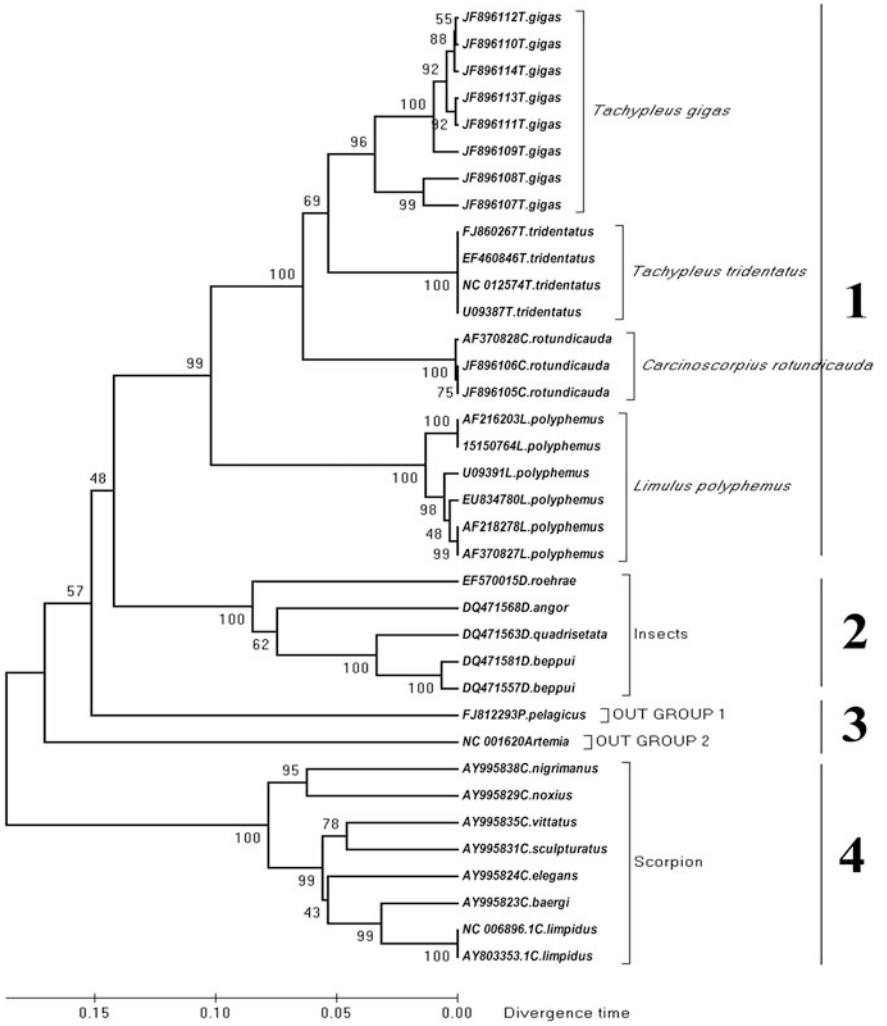

Fig. 2 Un-weighted Pair Group Method with Arithmetic Mean (UPGMA) phylogenetic tree was constructed to determine the evolutionary history of horseshoe crabs with related sister taxa (Saitou and Nei 1987)

file used for NJ method was also used for phylogram construction. As observed in NJ method, the out groups used were clearly segregated in separate branch in phylogram proving its reliability besides the higher bootstrap values in the internal branch nods. Four distinct clads were noted in the phylogram having horseshoe crabs in Clad 1, Insects in Clad 2, out group organisms in Clad 3 and Scorpions in Clad 4. Unlike the NJ method, UPGMA method could not show more apparently the phylogeographical signals in the constructed phylogram. However, similar to

NJ method, UPGMA method also proved the efficiency of cox1 gene in species level identification of various developmental stage of T. gigas but it was inefficient in sex determination and precise identification of various developmental stages. The constructed phylogram also proved the monophyletic nature of T. gigas with T. tridentatus and their closer genetic relatedness. Atlantic horseshoe crab (L. polyphemus) was genetically distinct from other species of horseshoe crabs. Insects were clumped together in Clad 2 which shared a common branch node with horseshoe crabs proving their closer genetic relatedness with horseshoe crabs. As observed in NJ method, UPGMA method also segregated terrestrial scorpion species into a separate clad (Clad 4). Depth analysis of horseshoe crab phylogene using cox1 gene as a reference sequence clearly showed the closer genetic relatedness of *T. gigas* with T. tridentatus and their monophyletic origin which other molecular markers failed to address (Fig. 3).

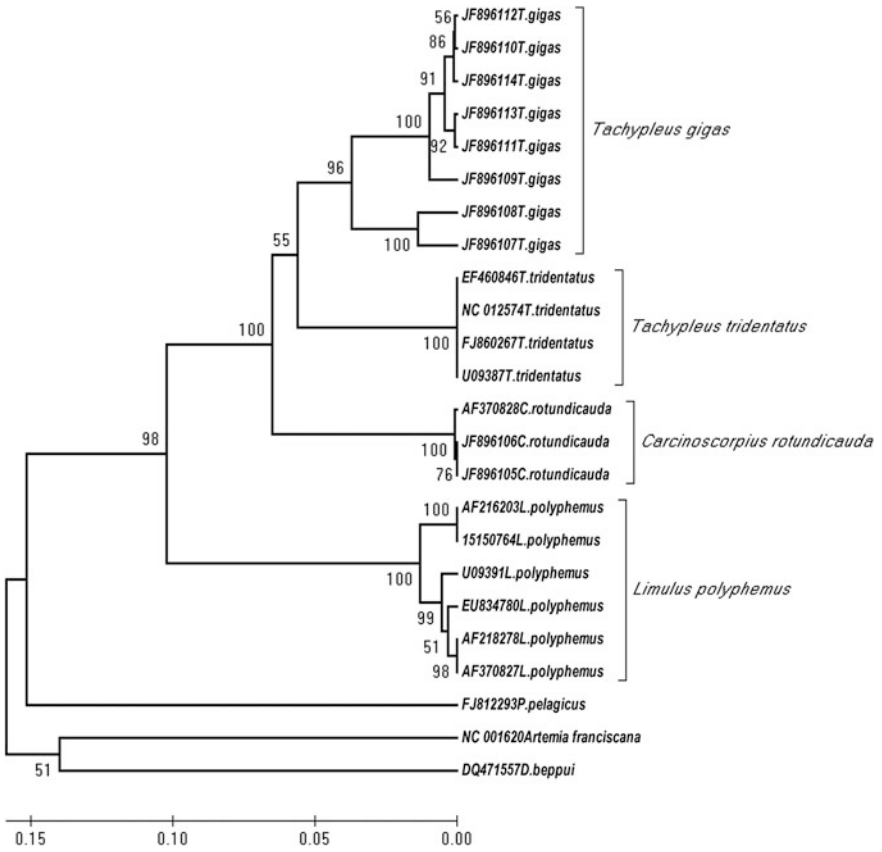

Fig. 3 Monophyletic origin of *Tachypleus gigas* with *T. tridentatus* was determined using UPGMA phylogenetic tree

Genetic Distance Data Analysis

Mean genetic distance within *L. polyphemus* at 1st, 2nd and 3rd codon position was 0.006, 0 and 0.048 respectively. Intra-species genetic distance among *T. gigas* and *C. rotundicauda* at 1st, 2nd and 3rd codon position were 0.016, 0.03 and 0.064 (*T. gigas*) and 0.003, 0 and 0.05 (*C. rotundicauda*) respectively. *T. tridentatus* showed no variation in GD value in all the codon positions. Among the Asian horseshoe crab species *T. gigas* showed closer genetic relatedness (lower GD value) with *T. tridentatus* with GD values of 0.037, 0.017 and 0.307 at 1st, 2nd and 3rd codon position respectively. This observation clearly revealed their monophyletic origin. The genetic distance values between *C. rotundicauda* and *T. gigas* were 0.053, 0.017 and 0.395 at 1st, 2nd and 3rd codon position respectively. Calculated genetic distance data showed higher genetic distance value in third codon position than its corresponding first and second codon positions. The GD data also proved the distance of genetic relatedness of *L. polyphemus* with Asian horseshoe crabs (Table 3).

Mean GD value within the four species of horseshoe crabs was 0.132, 0.053, 0.019 and 0.405 at 1st, 2nd and 3rd codon position respectively (at 1st + 2nd + 3rd + Noncoding gene). Average GD value between representative Insects and horseshoe crabs at 1st, 2nd and 3rd codon position were 0.207, 0.077 and 0.718 respectively. On the other hand mean GD values between representative

Table 3 Average genetic distance (GD) between four available species of horseshoe crabs observed in only 1st, 2nd and 3rd codon positions

		L. polyphemus			*T. gigas*			*T. tridentatus*			*C. rotundicauda*		
	CP	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
LP	1st	0.006			0.092			0.077			0.080		
	2nd		0			0.032			0.015			0.015	
	3rd			0.048			0.659			0.66			0.629
TG	1st				0.016			0.037			0.053		
	2nd					0.030			0.017			0.017	
	3rd						0.064			0.307			0.395
TT	1st							0			0.037		
	2nd								0			0	
	3rd									0			0.363
CR	1st										0.003		
	2nd											0	
	3rd												0.05

Note CP—Codon positions; LP—*L. polyphemus*; TG—*T. gigas*; TT—*T. tridentatus*; CR—*C. rotundicauda*

Scorpion and horseshoe crabs at 1st, 2nd and 3rd codon position were 0.217, 0.063 and 1.264 respectively (Table 4). These observations proved the closer genetic relatedness of horseshoe crabs with insects than with scorpions.

Nucleotide diversity between horseshoe crab species was comparatively smaller ($\pi = 0.150633$) than between related sister group taxa and horseshoe crabs. Positive Tajima test statistic values of overall and within horseshoe crabs population signified low levels of both low and high frequency polymorphisms in the sequences (Table 5).

Nucleotide substitution pattern observed among the test organisms (including horseshoe crabs, insects, scorpions and common crabs) clearly showed that transitional substitutions are very common in the gene sequence than transversional substitutions with transition/transversion bias value $R = 0.988$. The transition/transversion rate ratios were $k_1 = 2.243$ (for purines) and $k_2 = 2.414$ (for pyrimidines) (Table 6).

Nucleotide substitution pattern observed within the horseshoe crabs also showed that transitional substitutions are very common in the gene sequence than transversional substitutions with transition/transversion bias value $R = 1.711$. The transition/transversion rate ratios were $k_1 = 4.143$ (for purines) and $k_2 = 4.375$ (for pyrimidines) (Table 7).

Table 4 Mean genetic distance (GD) values of different groups of organisms with reference to horseshoe crab at all the possible codon position indicating 3rd codon position shows higher GD value

	1st + 2nd + 3rd codon position	1st codon position	2nd codon position	3rd codon position
	Horseshoe crabs	Horseshoe crabs	Horseshoe crabs	Horseshoe crabs
Horseshoe crabs	0.132	0.053	0.019	0.405
Insect	0.284	0.207	0.077	0.718
Scorpions	0.377	0.217	0.063	1.264

Table 5 Results from Tajima's Neutrality test calculated for 36 sequences (including horseshoe crabs, insects and scorpions) and 24 sequences (only horseshoe crabs)

	m	S	p_s	π	D
Overall	36	332	0.536349	0.210329	2.372310
Between horseshoe crabs	24	290	0.461783	0.150633	0.878823

Note The Tajima test statistic (Tajima 1989) was estimated using MEGA4 (Tamura et al. 2007). All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). The abbreviations used are as follows: m = number of sites, S = Number of segregating sites, $p_s = S/m$, $\Theta = p_s/a_1$, and π = nucleotide diversity. D is the Tajima test statistic

Table 6 Pattern of Nucleotide substitution observed among the test organisms (including horseshoe crabs, insects, scorpions and common crabs) calculated using Maximum Composite Likelihood method

	A	T	C	G
A	–	*8.43*	*4.35*	**9.79**
T	*5.91*	–	**10.49**	*4.37*
C	*5.91*	**20.36**	–	*4.37*
G	**13.25**	*8.43*	*4.35*	–

Note Each entry shows the probability of substitution from one base (row) to another base (column) instantaneously. Only entries within a row should be compared. Rates of different transitional substitutions are shown in bold and those of transversional substitutions are shown in italics. The nucleotide frequencies are 0.256 (A), 0.366 (T), 0.189 (C), and 0.189 (G). The overall transition/transversion bias is R = 0.988, where $R = \frac{A \times G \times k1 + T \times C \times k}{(A+G) \times (T+C)}$. Codon positions included were 1st + 2nd + 3rd + Noncoding. All positions containing gaps and missing data were eliminated from the dataset (Complete-deletion option)

Table 7 Pattern of Nucleotide substitution observed within horseshoe crabs, calculated using Maximum Composite Likelihood method

	A	T	C	G
A	–	*5.47*	*3.33*	**11.05**
T	*4.49*	–	**14.56**	*2.67*
C	*4.49*	**23.91**	–	*2.67*
G	**18.59**	*5.47*	*3.33*	–

Note Each entry shows the probability of substitution from one base (row) to another base (column) instantaneously. Only entries within a row should be compared. Rates of different transitional substitutions are shown in bold and those of transversional substitutions are shown in italics. The nucleotide frequencies are 0.281 (A), 0.343 (T), 0.209 (C), and 0.167 (G). The overall transition/transversion bias is R = 1.711, where $R = \frac{A \times G \times k1 + T \times C \times k}{(A+G) \times (T+C)}$. Codon positions included were 1st + 2nd + 3rd + Noncoding. All positions containing gaps and missing data were eliminated from the dataset (Complete-deletion option)

4 Discussion

4.1 Phylogenetic Study on Horseshoe Crab

Earlier studies on the serology of Atlantic species, *(L. polyphemus)* and Indo-Pacific conspecifics (Shuster 1962), phylogenetic analyses of amino acid sequences of coagulogen and the fibrinopeptide-like peptide C (Shishikura et al. 1982; Srimal et al. 1985; Sugita and Shishikura 1995), immunological comparisons of hemocyanins (Sugita 1988), two-dimensional electrophoresis of general proteins (Miyazaki et al. 1987), interspecific hybridization experiments (Sekiguchi and Sugita 1980), cladistic appraisals of morphological characters (Fisher 1984), and

mtDNA genes (Avise 1994) had clearly proved the sister taxon status of *L. polyphemus* to Indo-Pacific conspecifics. However, these molecular tools failed to address the monophyletic origin of *Tachypleus gigas* with *T. tridentatus.* Instead, all these molecular tools clumped *T. gigas* with *C. rotundicauda* and showed lower genetic distance between *T. gigas* and *C. rotundicauda* than between *T. gigas* and *T. tridentatus* (Kamaruzzaman et al. 2011). It was strongly believed that these three species constitute a phylogenetically irresolvable trichotomy (Xia 2000). However, all the previous studies failed to address the sampling size used in the phylogenetic tree construction. Ward et al. (2005) proved the importance of sampling size in determining and identifying the species using inter and intra species genetic distance data analysis. Recent studies on horseshoe crab phylogeny (Xia 2000; Kamaruzzaman et al. 2011) have a serious drawback due to the sample size they used to evaluate the topology of the phylogenetic tree. In this study, considerable number of sample (only horseshoe crabs N = 21) size gave substantial amount of information on the monophyletic origin of *T. gigas* and *T. tridentatus* which in turn was verified using UPGMA phylogram and genetic distance data analysis. Present study also showed the phylogeographical cues in cox1 gene.

4.2 Distance Matrix Method

4.2.1 Neighbor Joining (NJ) Tree

The phylogeographical signals were apparent in Clad 1, segregating *T. gigas* collected from Pulau Gaya (East Malaysia: Borneo) from the very species sampled in East coast of Peninsular Malaysia (Balok and Pekan). But the phylogram failed to segregate the samples collected from Balok and Pekan. This might probably due to 1. The short geographical distance between these sampled area (~80 km), 2. Constant gene flow between the horseshoe crab populations from these two sampling stations. Similar observation was reported by various researchers on fishes (Akbar John et al. 2010), ticks (Song et al. 2011) spiders (Zhang et al. 2005) and others (Hickerson and Cunningham 2000; MuÑOz et al. 2008). As observed in previous studies, the distant relatedness of American horseshoe crab (*L. polyphemus*) to the Indo Pacific conspecifics was apparent in the phylograms. More interestingly, the monophyletic origin of *T. gigas* with *T. tridentatus* was evident in Neighbor Joining (NJ) phylogram and this observation was cross examined using genetic distance (GD) data analysis which showed the GD value of (0.037, 0.017 and 0.307 at 1st, 2nd and 3rd codon position respectively) between *T. gigas* and *T. tridentatus*, whereas the GD values between *T. gigas* and *C. rotundicauda* was 0.053, 0.017 and 0.395 at 1st, 2nd and 3rd codon position respectively. Though this observation is in contrast to the previous studies, it is virtually true because of the large number of individual species of horseshoe crab samples used in this study. The monophyletic origin of *T. gigas* and *T. tridentatus* and the reliability of the phylogram were also checked by constructing sub phylogram constituting only

representative horseshoe crab species and the selected out groups from possible sister taxa which also showed the similar results.

The phylogram clearly showed the genetic relatedness of horseshoe crabs and insects indicating that horseshoe crabs might have probably evolved from the ancient aquatic insects. This observation is in agreement with the conclusions of recent studies on horseshoe crab phylogeny (Xia 2000; Kamaruzzaman et al. 2011). However, Eurypterids (e.g., sea scorpions) have traditionally been regarded as close relatives of horseshoe crabs. Subsequent studies placed eurypterids closer to the arachnids (e.g., spiders, terrestrial scorpions, mites and ticks) in a group called Metastomata (Pavlicek et al. 2008). There has also been a belief that eurypterids are closely related to terrestrial scorpions (Raz et al. 2009). Recent study on the genetic relationships between arachnids and their relatives recognized Eurypterida, Xiphosura and Arachnida as three major groups, and their genetic relatedness at present cannot be resolved using available molecular markers (Shultz 2007). Similar conclusion was evident in the phylogram which was cross checked with genetic distance data that terrestrial scorpions which are closely related to sea scorpions are distantly related to horseshoe crabs.

4.2.2 PGMA Method

Unweighted Pair Group Method with Arithmetic Mean (UPGMA) was used to construct un-rooted phylogenetic tree to 1. Infer the genetic relationship between horseshoe crabs, 2. To check the reliability of information on the monophyletic origin of *T. gigas* with *T. tridentatus,* and 3. To verify the observation obtained in NJ tree method. As noted in NJ tree, UPGMA tree had 4 distinct clads segregating representative species to their respective clads. Unlike NJ tree, the basic principle and hypothesis in UPGMA method (constant rate of evolution) restricted its efficiency in showing the distinct phylogeographical cues in the phylogram. However, higher bootstrap value in internal nods and distinct segregation of representative species showed its efficiency in phylogenetic tree construction. The phylogram also proved the monophyletic origin of *T. gigas* with *T. tridentatus* and their closer genetic relatedness. Atlantic horseshoe crab (*L. polyphemus*) was genetically distinct from other species of horseshoe crabs. To verify this observation, a sub tree was constructed consisting of only 4 species of horseshoe crab and 3 out group species which also showed similar facts. Insects were clumped together in Clad 2 which shared a common branch node with horseshoe crabs proving their closer genetic relatedness with horseshoe crabs. Phylogram also showed the distant genetic relatedness of terrestrial scorpions with horseshoe crabs similar to NJ tree. Both NJ and UPGMA phylogram clearly proved the efficacy of COI gene in delineating the members of evolutionarily cryptic groups of organisms, besides revealing the monophyletic origin of south East Asian horseshoe crabs (*T. gigas* and *T. tridentatus*).

4.3 Genetic Distance (GD) Data Analysis

The basic principle behind the DNA barcoding technology is there should be low rates of DNA sequence divergence among individuals of the same species than between the species. In other words, intra species genetic distance should be low compared to inter species genetic distance. These were quite evident in GD data analysis (Hebert et al. 2003, 2004). In this study we mainly concentrated on the calculation of GD within horseshoe crabs because of the representation of all the four extant species in the constructed phylogram. Mean genetic distance within the horseshoe crabs at all the codon positions were 0.053, 0.019 and 0.405 at 1st, 2nd and 3rd Codon position respectively. Genetic Distance between the selected animal groups showed that the GD value was lower between horseshoe crabs and insects with 0.207, 0.077 and 0.718 at 1st, 2nd and 3rd codon position respectively compared to horseshoe crabs and scorpions with 0.217, 0.063 and 1.264 codon position respectively. This observation clearly indicated the closer genetic relatedness of insects with horseshoe crabs. Similar observation was made by Kamaruzzaman et al. (2011). It was also observed that the horseshoe crabs are genetically related to common crabs than the scorpion. This might be the reason why out group (*Portunus pelagicus*) used in this study clustered with horseshoe crabs closely than scorpions.

Another interesting observation made from the genetic distance data was higher genetic distance observed in third codon position than its corresponding first and second codon positions. Similar observation was made by Ward et al. (2005) while barcoding fishes from Australian waters. Simmons et al. (2006) also observed that greater phylogenetic signal is often found in parsimony-based analyses of third codon positions of protein-coding genes relative to their corresponding first and second codon positions, even for early-derived basal clades (Siemion and Przemyslaw 1994; Ajmal Khan et al. 2010). Average genetic distance among the different groups of test organisms used in this study showed higher GD value at 3rd codon position indicating that detailed study on 3rd codon position might reveal possible evolutionary information among the closely related groups of organisms.

4.4 Tajima's Neutrality Test

Nucleotide diversity between horseshoe crab species was comparatively smaller ($\pi = 0.150633$) than between related sister group taxa and horseshoe crabs ($\pi = 0.210329$). Positive Tajima test statistic values of overall and within horseshoe crabs population signified that there were low levels of both low and high frequency polymorphisms in the horseshoe crab sequences which ultimately helped them in retaining the genetic makeup virtually unchanged over millions of years.

4.5 Nucleotide Substitution Analysis

It is a well-known fact that during DNA sequence evolution the rate of transitional changes differs from the rate of transversional changes, with transitions generally occurring more frequently than transversions. This difference is often referred to as transition bias, and estimation of the extent of transition bias may be of interest, since it may vary for different organisms and for different genes within a collection of organisms. In general, there are twice as many possible transversions as transitions due to the relatively high rate of mutation of methylated cytosines to thymine (Brown et al. 1982; Gojobori et al. 1982; Curtis and Clegg 1984; Graur and Li 2000). Proper estimation is also important because the ratio of the rates of transitional to transversional changes (often called the Ti:Tv ratio) play a role in evolutionary distance correction methods and is used in several common evolutionary models (e.g., the F84 model) (Wakeley 1996). However, it was observed that Ti:Tv ratio is strongly influenced by sampling size. Hence, we also calculated ti/tv bias which is more realistic and widely applied reliable estimate. Transition/Transversion (ti/tv) bias is known to be a general property of DNA sequence evolution, it is more pronounced in animal mitochondrial DNAs (mtDNAs) than in nuclear or chloroplast DNAs (Wakeley 1996). Estimation of the ti/tv bias is important not only to our understanding of the patterns of DNA sequence evolution, but also to reliable estimation of sequence distance and phylogeny reconstruction (Rosenberg et al. 2003).

The calculated ti/tv bias among the test organisms (including horseshoe crabs, insects, scorpions and common crabs) was $R = 0.988$ whereas ti/tv bias within horseshoe crabs was 1.711. This observation clearly proved that horseshoe crab *cox1* undergoes more transition mutations compared to transversion mutations. This observation was also proved by calculating Ti:Tv ratio where this ratio was $k_1 = 4.143$ (for purines) and $k_2 = 4.375$ (for pyrimidines) within the horseshoe crabs.

5 Conclusion

Molecular taxonomic study clearly segregated American conspecific (*Limulus polyphemus*) from Indo Pacific horseshoe crabs and thereby proving their distant intra species genetic relatedness. More interestingly, the efficiency of *cox1* gene in species level delineation of the cryptic taxonomy of horseshoe crabs proved the monophyletic origin of *Tachypleus gigas* and *T. tridentatus* which other molecular markers failed to address. Close genetic relatedness of horseshoe crabs with insects showed that they might have evolved from ancient aquatic insects. The efficiency of cytochrome oxydase C subunit 1 gene in species level identification evolutionarily conserved horseshoe crab genome was apparent in the constructed phylogram together with the precise identification of their differential developmental stages to the species level. It was also evident from the phylogram that *cox1* gene has sound

phylogeographical signals. Higher Genetic Distance (GD) value obtained from 3rd codon position than its corresponding 1st and 2nd codon positions proved the presence of grater genetic cues in 3rd codon position that could be used to study the genetic relatedness of evolutionarily conserved species in future.

References

Ajmal Khan S, Lyla PS, Akbar John B, Prasanna Kumar C, Murugan S, Jalal KCA (2010) DNA barcoding of Stolephorus indicus, Stolephorus commersonnii and Terapon jarbua of Parangipettai coastal waters. Biotechnology 9:373–377

Akbar John B, Kamaruzzaman BY, Jalal KCA, Zaleha K (2011) Hydrology of the horseshoe crab nesting grounds at Pahang coast, Malaysia. Orient J Chem 27(4):1475–1483

Akbar John B, Kamaruzzaman BY, Jalal KCA, Zaleha K (2012) Sediment profiling of the nesting grounds of horseshoe crabs at east peninsular Malaysia. Int J Biol 4(2):159–165

Akbar John B, Prasanna kumar C, Lyla PS, Ajmal Khan S, Jalal KCA (2010) DNA barcoding of lates calcarifer (Bloch 1970). Res J Biol Sci 5:414–419

Avise JC (1994) Molecular markers, natural history and evolution. Chapman & Hall, New York

Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences of primates: tempo and mode of evolution. J Mol Evol 18(4):225–239

Chiu HMC, Morton B (2003) The morphological differentiation of two horseshoe crab species, Tachypleus tridentatus and Carcinoscorpius rotundicauda (Xiphosura), in Hong Kong with a regional Asian comparison. J Nat Hist 37(19):2369–2382

Curtis SE, Clegg MT (1984) Molecular evolution of chloroplast DNA sequences. Mol Biol Evol 1 (4):291–301

Fisher DC (1984) The Xiphosurida: archetypes of bradytely. In: Eldredge N, Stanley SM (eds) Living fossils. Springer, New York, pp 196–213

Gojobori T, Li WH, Graur D (1982) Patterns of nucleotide substitution in pseudogenes and functional genes. J Mol Evol 18(5):360–369

Graur D, Li WH (2000) Fundamentals of molecular evolution. Sinauer, Sunderland

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proc R Soc Lond Ser B Biol Sci 270(1512):313–321

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. Proc Natl Acad Sci USA 101(41):14812–14817

Hickerson MJ, Cunningham CW (2000) Dramatic mitochondrial gene rearrangements in the hermit crab Pagurus longicarpus (Crustacea, Anomura). Mol Biol Evol 17(4):639–644

Kamaruzzaman BY, John BA, Zaleha K, Jalal KCA (2011) Molecular phylogeny of horseshoe crab. Asian J Biotechnol 3(3):302–309

Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16:111–120

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23(21):2947–2948

Mikkelsen T (1988) The secret in the blue blood. Science Press, Beijing

Miyazaki JI, Sekiguchi K, Hirabayashi T (1987) Application of an improved method of two-dimensional electrophoresis to the systematic study of horseshoe crabs. Biol Bull 172 (2):212–224

MuÑOz J, GÓMez A, Green AJ, Figuerola J, Amat F, Rico C (2008) Phylogeography and local endemism of the native Mediterranean brine shrimp Artemia salina (Branchiopoda: Anostraca). Mol Ecol 17(13):3160–3177

Obst M, Faurby S, Bussarawit S, Funch P (2012) Molecular phylogeny of extant horseshoe crabs (Xiphosura, Limulidae) indicates Paleogene diversification of Asian species. Mol Phylogeny Evol 62(1):21–26

Pavlicek T, Mienis HK, Raz S, Hassid V, Rubenyan A, Nevo E (2008) Gastropod biodiversity at the evolution canyon microsite, lower Nahal Oren, Mount Carmel, Israel. Biol J Linn Soc 93:147–155

Prasanna Kumar C, Akbar John B, Ajmal Khan S, Lyla PS, Murugan S, Rozihan M, Jalal KCA (2011) Efficiency of universal barcode gene (Coxi) on morphologically cryptic mugilidae fishes delineation. Trends Appl Sci Res 6:1028–1036

Raz S, Retzkin S, Pavlíček T, Hoffman A, Kimchi H, Zehavi D, Nevo E (2009) Scorpion biodiversity and interslope divergence at "evolution canyon", lower Nahal Oren microsite, Mt. Carmel, Israel. PLoS ONE 4(4):e5214

Rosenberg MS, Subramanian S, Kumar S (2003) Patterns of transitional mutation biases within and among mammalian genomes. Mol Biol Evol 20(6):988–993

Rudkin DM, Young GA (2009) Horseshoe crabs—an ancient ancestry revealed. In: Tanacredi JT, Botton ML, Smith DR (eds) Biology and conservation of horseshoe crabs. Springer, US, pp 25–44

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evo 4:406–425

Sekiguchi K, Nakamura K (1979) Ecology of the extant horseshoe crabs. In: Cohen E (ed) Biomedical applications of the horseshoe crab (Limulidae). Alan R. Liss, New York, pp 37–45

Sekiguchi K, Sugita H (1980) Systematics and hybridization in the four living species of horseshoe crabs. Evolution 34(4):712–718

Shishikura F, Nakamura S, Takahashi K, Sekiguchi K (1982) Horseshoe crab phylogeny based on amino acid sequences of the fibrino-peptide-like peptide C. J Exp Zool 223(1):89–91

Shultz JW (2007) A phylogenetic analysis of the arachnid orders based on morphological characters. Zoolog J Linn Soc 150:221–265

Shuster CN (1962) Serological correspondence among horseshoe crabs (Limulidae). Zoologica 47:1–8

Siemion IZ, Przemyslaw JS (1994) The informational context of the third base in amino acid codons. Biosystems 33:139–148

Simmons MP, Zhang L-B, Webb CT, Reeves A (2006) How can third codon positions outperform first and second codon positions in phylogenetic inference? An empirical example from the seed plants. Syst Biol 55(2):245–258

Song S, Shao R, Atwell R, Barker S, Vankan D (2011) Phylogenetic and phylogeographic relationships in Ixodes holocyclus and Ixodes cornuatus (Acari: Ixodidae) inferred from COX1 and ITS2 sequences. Int J Parasitol 41(8):871–880

Srimal S, Miyata T, Kawabata S-I, Miyata T, Iwanaga S (1985) The complete amino acid sequence of coagulogen isolated from Southeast Asian horseshoe crab, Carcinoscorpius rotundicauda. J Biochem 98(2):305–318

Sugita H (1988) Immunological comparisons of hemocyanins and their phylogenetic implications. In: Sekiguchi K (ed) Biology of horseshoe crabs. Science House, Tokyo, pp 315–334

Sugita H, Shishikura F (1995) A case of orthologous sequences of hemocyanin subunits for an evolutionary study of horseshoe crabs: amino acid sequence comparison of immunologically identical subunits of Carcinoscorpius rotundicauda and Tachypleus tridentatus. Zoolog Sci 12 (5):661–667

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24(8):1596–1599

Tajima F (1989) Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. Genetics 123:585–595

Wakeley J (1996) The excess of transitions among nucleotide substitutions: new methods of estimating transition bias underscore its significance. Trends Ecol Evol 11:158–163

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN (2005) DNA barcoding Australia's fish species. Philos Trans R Soc B Biol Sci 360(1462):1847–1857

Xia X (2000) Phylogenetic relationship among horseshoe crab species: effect of substitution models on phylogenetic analyses. Syst Biol 49(1):87–100

Yamasaki T, Makioka T, Saito J (1988) Morphology. In: Sekiguchi K (ed) Bioligy of horseshoe crabs. Tokyo, Science House Co. Ltd., pp 69–132

Zaleha K, Kamaruzzaman BY, Akbar John B, Ong MC (2010) Cd, Cu and Pb concentration levels in horseshoe crab nesting grounds of Pahang coast, Malaysia. J Biol Sci 10(8):790–794

Zhang AB, Kubota K, Takami Y, Kim JL, Kim JK, Sota T (2005) Species status and phylogeography of two closely related Coptolabrus species (Coleoptera: Carabidae) in South Korea inferred from mitochondrial and nuclear gene sequences. Mol Ecol 14(12):3823–3841