

Mitogenomic variation of Bangladesh *Penaeus monodon* (Decapoda, Penaeidae) and reassessment of its phylogeography in the Indo-West Pacific region

M. M. Mahbub Alam · Kristen M. Westfall ·
Snæbjörn Pálsson

Received: 14 April 2015 / Revised: 18 June 2015 / Accepted: 20 June 2015 / Published online: 29 June 2015
© Springer International Publishing Switzerland 2015

Abstract Large variation in the mtDNA control region (CR) in *Penaeus monodon* has been reported from the Indo-West Pacific region, characterizing three distinct lineages, A, B and C. This diversification has been explained by geographic barriers but may also be due to paralogous duplication events. Here we evaluate the diversity, origin and population patterns of this species in Bangladesh based on mtDNA, and reconsider the evidence for paralogous duplication. MtDNA variation revealed high diversity with almost all individuals carrying a unique haplotype, indicating a large effective population size and a single reproductive pool. The Bangladesh haplotypes cluster within lineage C, found mainly west of the Sunda Shelf. This lineage, which diverged from lineages A

and B around 6.6 Mya, has diverged further into several lineages in the western Indian Ocean. Sequences of mitochondrial CR, 16S rRNA gene and a combined fragment of the 16S rRNA and tRNA^{Val} genes from Bangladesh did not reveal any duplication or heteroplasmy. Further, phylogenies based on the three markers, including samples from different areas in the Indo-West Pacific region, support the occurrence of three distinct lineages. Thus, the variation in CR appears to result from segregation of orthologous variants rather than diversification following a paralogous duplication.

Keywords *Penaeus monodon* · Mitochondrial control region · Phylogeography · Paternal inheritance · Bangladesh · Indo-West Pacific

Electronic supplementary material The online version of this article (doi:10.1007/s10750-015-2381-3) contains supplementary material, which is available to authorized users.

Handling editor: Diego Fontaneto

M. M. MahbubAlam (✉) · K. M. Westfall · S. Pálsson
Department of Life and Environmental Sciences,
University of Iceland, Reykjavik, Iceland
e-mail: mma3@hi.is

M. M. MahbubAlam
Faculty of Fisheries, Sylhet Agricultural University,
Sylhet, Bangladesh

K. M. Westfall
VÖR Marine Research Centre, Ólafsvík, Iceland

Introduction

Penaeus monodon (Fabricius, 1798), the giant tiger prawn, is widely distributed throughout the Indo-West Pacific region and is the most important penaeid shrimp for both aquaculture and capture fisheries (Motoh, 1985; Kumar et al., 2007; You et al., 2008; Mandal et al., 2012; Waqairatu et al., 2012; Vaseeharan et al., 2013). Due to its high nutrition value and economic importance, this species has been studied for sustainable aquaculture production (Kumar et al., 2007; Mandal et al., 2012) and has been a subject of numerous studies on genetic variation in order to

maintain its genetic diversity through management of wild populations (e.g. Benzie et al., 2002; You et al., 2008; Mandal et al., 2012; Waqairatu et al., 2012), but the sampling effort has been uneven in different geographic regions (Keyse et al., 2014).

Several marine species in the Indo-West Pacific region are characterized by distinct eastern and western genetic lineages (Gaither et al., 2001; Keyse et al., 2014), including *P. monodon* (Benzie et al., 2002) and *Fenneropenaeus indicus* (Milne-Edwards, 1837; Alam et al., 2015). These phylogeographic patterns may have resulted from restricted genetic connectivity of populations across the Sunda Shelf during periods of low sea level in the glacial periods of the Pleistocene (Voris, 2000). In *F. indicus* distinct lineages were also found west of the Sunda Shelf, pointing to geographic isolation of the south-east African coast from the Bay of Bengal (Alam et al., 2015), in accordance to the split of the Western Indian Province from other regions in the Indo-West Pacific region (Bowen et al., 2012).

Penaeus monodon is known for high variation in molecular markers, which is geographically structured in the Indo-West Pacific region (Benzie et al., 2002; Kumar et al., 2007; You et al., 2008; Zhou et al., 2009; Mandal et al., 2012; Waqairatu et al., 2012). Variation in the mitochondrial DNA has been widely assessed for taxonomic identification and in phylogeographic and population genetic studies e.g. control region (CR) (Walther et al., 2011; Vaseeharan et al., 2013), and the 16S rRNA gene and a downstream fragment of the 16S rRNA and tRNA^{Val} genes (Lavery et al., 2004; Kumar et al., 2007; Pascoal et al., 2008; Vaseeharan et al., 2013). The highest mtDNA diversity in *P. monodon* was found in Indonesia, with less diversity identified in south-east Africa and west Australia (Benzie et al., 2002). High genetic variation was also reported in the Indian population, being highest in the Andaman region followed by the east coast and the west coast (Kumar et al., 2007; Mandal et al., 2012). *Penaeus monodon* populations in the Indo-West Pacific region were broadly structured as three mtDNA groups, i.e. south-east African, south and south-east Asian and Pacific populations (Benzie et al., 2002; You et al., 2008; Waqairatu et al., 2012). The south and south-east Asian group is further divided into two groups, west and east of the Sunda Shelf (Benzie et al., 2002; Kumar et al., 2007; You et al., 2008; Mandal et al., 2012). The populations in the south-east African group

were unique and genetically differentiated from the other main groups, which were partially admixed (Kumar et al., 2007; You et al., 2008; Mandal et al., 2012). The high genetic diversity in Indonesia was explained by Benzie et al. (2002), which was resulted from admixture due to migration of the divergent African populations into south-east Asia.

The high mtDNA diversity in *P. monodon* was proposed by Walther et al. (2011) to result from a paralogous duplication event but studies of the mitochondrial CR defined three distinct lineages across the species distribution (A, B and C); A was mostly found in the Pacific Ocean region, B from the juncture of the Indian-West Pacific Oceans and C from India and Africa (Zhou et al., 2009; Walther et al., 2011). Walther et al. (2011) detected different haplotypes within a subset of individuals from Indonesia and concluded that these could have arisen due to duplication, either due to insertions of mitochondrial sequences into the nuclear genome (numts) or within the mitochondrial genome, and concluded that previous studies had thus highly overestimated the mtDNA variation. Although Walther et al. (2011) acknowledged that the occurrence of such divergent lineages could be due to heteroplasmy as a result of bi-parental inheritance as has been observed in crosses between genetically distinct individuals (Dokianakis & Ladoukakis, 2014), it did not affect their main conclusion.

Nuclear genetic variation in *P. monodon*, as assessed by allozymes, DNA sequences and microsatellites, shows patterns at both small and large geographic scales. In allozymes, differences between the north and west coasts of Australia reflect colonization by populations from the east (Benzie et al., 1992, 1993). In contrast, South Africa, Mozambique and Madagascar, show panmixia with high gene flow among all populations (Forbes et al., 1999). Microsatellites revealed a clear structure in the Indo-West Pacific populations of *P. monodon* (You et al., 2008; Waqairatu et al., 2012) and a split between the eastern Indian Ocean and western Pacific, based on the elongation factor 1-alpha gene, was observed by Duda & Palumbi (1999).

Penaeus monodon is the most economically targeted shrimp species in Bangladesh (Quader, 2010), but despite this, no molecular studies have been conducted on this species from Bangladesh, and only a single phylogeographic study has included a Penaeid shrimp from the country (Alam et al., 2015). The

diversity of shrimps is under continuous threat due to post-larvae collection, coastal and marine pollution, impacts of coastal aquaculture, natural disasters, sea level rise, persistent organic pollutants, etc. However, the population of *P. monodon* appears to be large, considering the annual harvest sustained by suitable habitats in Mangrove forests and large rivers (Quader, 2010; Department of Fisheries Bangladesh, 2013). Therefore, a study on population structure and genetic diversity of *P. monodon* in Bangladesh waters and in comparison to other Indo-West Pacific regions is of great importance for proper management and conservation of the species.

The main objectives of this study are to analyse the genetic diversity and origin of *P. monodon* in Bangladesh, based on mtDNA variation, and to reevaluate the larger phylogeographic patterns within the Indo-West Pacific region. To reach our objectives we hypothesize that (1) there is a high diversity within the Bangladesh population, as had been previously reported in other Indo-West Pacific regions, reflecting a large and single panmictic population which has expanded after the last glacial period of the Ice Age, (2) the Bangladesh population shares its most recent ancestry with the other populations from the Bay of Bengal (India, Sri Lanka and Thailand) and the distinct lineages within the Indo-West Pacific region correspond to the major geographic barriers in the region and (3) the sequence variation within mitochondrial CR reflects orthologous variation.

Materials and methods

Sample collection and DNA extraction

A total of 86 wild origin *P. monodon* were collected from four locations along the Bangladesh coastline (Fig. 1) during the period of December 2012 to September 2013 and preserved in 96% ethanol. Artisanal fishermen collected 17 sub-adult specimens (average weight of 40 g) from Sundarban mangrove forest (SB). Samples of adults, with an average weight of 100 g, were collected from the Barguna coast (BC, 10 individuals), Middle ground (one of the four fishing grounds) of the Bay of Bengal (MB, 32 individuals) and the coast of Sent Martin's Island (SM, 27 individuals) by fish/shrimp trawlers. In addition, five specimens of *P. monodon* of wild origin were sampled

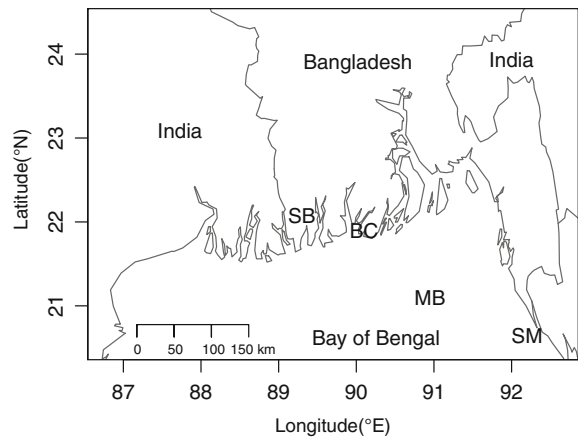


Fig. 1 Sampling sites along the coast of Bangladesh. Capital letters indicate the four sampling sites- SB Sundarban, Sathkhira, BC Barguna coast, Barguna, MB Middle ground (one of the four fishing grounds), Bay of Bengal and SM Sent Martin's Island, Cox's Bazar

from Negombo Lagoon ($n = 3$) and Trinco ($n = 2$) of Sri Lanka in 2008 and 2005, respectively. Total genomic DNA was extracted from ~ 1 mg pleopod tissue through overnight incubation at 56°C in a mixture of 6% Chelex and 0.2 mg ml^{-1} proteinase K followed by 10 min at 95°C .

Polymerase chain reaction (PCR) and sequencing

PCR was performed in a final volume of $10\ \mu\text{L}$, which included 30–150 ng DNA, 0.2 mM dNTP, 0.1% Tween 20, $1 \times$ Standard Taq Buffer (New England Biolabs), 0.5 mg Bovine Serum Albumin, 0.5 U Taq Polymerase and 0.34 mM each of forward and reverse primers. Three regions of mtDNA were amplified and sequenced; the CR, a fragment of the 16S rRNA gene and a combined fragment of the 16S rRNA and tRNA^{Val} genes. A 939 bp fragment of the CR was obtained using one reverse primer 1R (Chu et al., 2003) and the forward primer, PmCON-2F (Wilson et al., 2000) from 28 specimens. To look for the existence of heteroplasmy or duplicated sequences, as done by Walther et al. (2011), and to confirm the sequences, a second forward primer, 12S (Chu et al., 2003) was used together with the same reverse primer (1R) to amplify and sequence a 517 bp fragment from 82 specimens, including 24 specimens which had been sequenced with PmCON-2F. The shorter fragment from 14 specimens was also resequenced with the reverse primer 1R. In total, 86 sequences were

obtained from Bangladesh (GenBank accession numbers: KT006166–KT006251), and a shared aligned region of 517 bp was used for the downstream analyses of genetic diversity and demographic changes. A 370 bp sequence of the 16S rRNA gene was obtained using 16STf (MacDonald et al., 2005) and 16Sbr (Palumbi et al., 1991) primers, for 10 (of which 4 specimens were sequenced from both directions) and five individuals from Bangladesh and Sri Lanka, respectively (GenBank accession numbers: KT006252–KT006266). A 491 bp sequence of the combined 16S rRNA and tRNA^{Val} region was amplified from both forward and reverse directions using 16ScruC4 and 16ScruC3 primers (Pascoal et al., 2008) for five and three individuals from Bangladesh and Sri Lanka, respectively (GenBank accession numbers: KT006158–KT006165). Information about the primers in this study are given in Appendix 1.

The amplification protocol of the CR fragments included an initial denaturation at 94°C for 4 min, 37 cycles of denaturation at 94°C for 30 s, annealing at 49°C for 45 s and extension at 72°C for 1 min, and a final extension at 72°C for 6 min. The protocols for the 16S rRNA gene and the combined fragment of 16S rRNA and tRNA^{Val} genes were altered by increasing the annealing temperature to 55°C. The PCR products were sequenced with the Big Dye Terminator kit 3.1 (Applied Biosystems) and run on a Genetic Analyser (3500xL Applied Biosystems). All sequences were edited and aligned using ClustalW Multiple alignment implemented in Bioedit Sequence Alignment Editor (Hall, 1999). Detailed information of the sequences, including all GenBank sequences utilized for this study, is shown in Table 1 and Appendix 2.

Genetic diversity of *P. monodon* within Bangladesh

Genetic diversity of Bangladesh *P. monodon*, based on 86 CR sequences (517 bps), was summarized for each sampling site and across all sites, by calculating gene diversity (h) and nucleotide diversity (π), using ARLEQUIN v3.5 (Excoffier & Lischer, 2011). Haplotype richness was calculated using the hiefstat package (Goudet, 2005) in R (R Core Team, 2014). The relationship among CR haplotypes was investigated by constructing an unrooted network using a median-joining algorithm in NETWORK version 4.6.1.3 (Bandelt et al., 1999). Analysis of molecular

variation (AMOVA) of the samples from Bangladesh was conducted based on pairwise differences among sequences, and the proportion of variation among groups (Φ_{ST}) tested with 1000 permutations, using ARLEQUIN v3.5 (Excoffier & Lischer, 2011). The total genetic diversity of the samples from Bangladesh was compared with the estimated genetic diversities of other populations in the Indo-West Pacific region, based on data listed in Table 1 and Appendix 2.

Population demographic changes in *P. monodon* within Bangladesh were estimated, based on the 86 CR sequences (517 bps), by comparing the fit of the distribution of pairwise nucleotide differences (mismatch distribution) with the expected values of a demographic expansion, following the method developed by Rogers & Harpending (1992). The fit was estimated by the sum of square deviation (SSD) and the raggedness index (Harpending, 1994), and tested using ARLEQUIN v3.5 (Excoffier & Lischer, 2011). The time since expansion was based on the median of the mismatch distribution (τ) and the mutation rate, $\mu = 3.44\%$ per site per Myr, for the CR (see below), as $t = \tau/(2\mu)$, where l is the length of the sequence. Tajima's D (1989a, b, 1993) and Fu's F (1997) tests were performed, to further estimate the demographic changes, or possible deviation from neutrality, using ARLEQUIN. In addition, the demographic changes were evaluated using the bayesian skyline plot (BSP) analysis using BEAST v1.7.5 (Drummond et al., 2007). The BSP analysis estimated the posterior probability of the effective population size (N_e) using Markov Chain Monte Carlo (MCMC) procedures by tracing the ancestry to the most recent common ancestor within *P. monodon*. Markov chains were run for 5.0×10^7 generations and sampled every 1000th generation. The general time reversible (GTR) model with gamma and invariant sites parameters (G+I) was used, derived from a PhyML Test (Guindon et al., 2010) using the APE package (Paradis, 2006) in R (R Core Team, 2014). A strict molecular clock of 3.44% per million years (Myr) was used for the BSP analysis, which was calibrated for the CR based on a comparison of the divergence between Penaeid species for the CR and COI region (see calibration below). Bayesian Skyline was used as prior for the coalescent analysis, with a piecewise-constant as a skyline model. Log files were visualized for the posterior distributions of the Markov Chain statistics using TRACER v1.5 (Rambaut & Drummond, 2009),

Table 1 Information about the sequences for different fragments/genes of the mitochondrial DNA utilized for the study

Fragment/gene	Sampling location	Sample size	Reference	
Control region (CR)	Bangladesh	86	This study	
	Sri Lanka	26	*1	
	India (West coast)	2	Kumar et al. (2007), *2	
	India (South coast)	23	Kumar et al. (2007), *3	
	India (East coast)	83	*2, 4	
	Thailand (West coast)	28	You et al. (2008)	
	Thailand (East coast)	32	You et al. (2008)	
	Philippines	36	You et al. (2008)	
	Vietnam	60	You et al. (2008), Oanh et al. (2010)	
	Indonesia	143	Walther et al. (2011)	
	Australia	29	You et al. (2008), Wilson et al. (2000)	
	China	129	You et al. (2008), Zhou et al. (2009)	
	Madagascar	25	You et al. (2008)	
	Kenya	12	You et al. (2008)	
		126	Mkare et al. (2014)	
	Total	840		
	16S rRNA	Bangladesh	10	This study
		Sri Lanka	5	This study
		India	2	Kumar et al. (2007)
Iran		1 [#]	*5	
Tunisia		1 [#]	*6	
Thailand		1 [#]	*7	
Malaysia		4	*8, 5	
China		23	Laveret al. (2004), *9, 10, 11	
Australia		2	Wilson et al. (2000)	
Fiji		1 [#]	Michel-Salzat & Bouchon (2000)	
Indo-West Pacific		1 [#]	Tong et al. (2000)	
Unknown		1 [#]	*12	
Total		52		
Combined 16S rRNA and tRNA ^{Val}		Bangladesh	5	This study
		Sri Lanka	3	This study
	Malaysia	4	Pascoal et al. (2008), (2012)	
	Indo-West Pacific	2	Pascoal et al. (2008), (2012)	

Table 1 continued

Fragment/gene	Sampling location	Sample size	Reference
	Australia	2	Wilson et al. (2000)
	Unknown	2	*13
	Total	18	

*1 D. H. N. Munasinghe, University of Ruhuna, Sri Lanka, Unpubl data

*2 M. S. Shekhar & G. Gopikrishna, Central Institute of Brackishwater Aquaculture, India, Unpubl data

*3 A. C. S. Reddy et al., Centre for Cellular and Molecular Biology, India, Unpubl data

*4 A. C. Reddy et al., Dr. Babasaheb Ambedkar Marathwada University Aurangabad, India, Unpubl data

*5 T. Jahromi et al., Persian Gulf and Oman Sea Ecological Research Institute, Iran, Unpubl data

*6 N. Besbes et al., Institut National des Sciences et Technologie de la Mer (INSTM), Tunisia, Unpubl data

*7 T. L. Chaicherdskul et al., Mahidol University, Thailand, Unpubl data

*8 S. H. Tan et al., Xiamen University, China, Unpubl data

*9 Z. Falin & J. Shigui, South China Sea Fisheries Research Institute, China, Unpubl data

*10 W. Mai & C. Hu, South China Sea Institute of Oceanology, China, Unpubl data

*11 F. Zhou and S. Jiang, South China Sea Fisheries Research Institute, China, Unpubl data

*12 M. Li, Ningbo University, China, Unpubl data

*13 A. Pascoal et al., University of Santiago de Compostela, Spain, Unpubl data

and 10% of the samples were discarded as burn-in during Skyline reconstruction. Skyline data was exported from TRACER v1.5 and redrawn in R (R Core Team, 2014).

The calibration of the molecular clock for the CR [3.44% (2.4–4.5) % per Myr] was based on the comparison of the divergence between Penaeid species for the CR and the COI region assuming a 2% (1.4–2.6) % sequence divergence per Myr as for the COI gene of the genus *Alpheus* (Knowlton et al., 1993; Knowlton & Weigt, 1998). The average dissimilarity between the CR region of *P. monodon* (EU426760.1) and *F. indicus* (FJ002577.1), *Fenneropenaeus chinensis* (Osbeck, 1765) (HM358499.1), *Marsupenaeus japonicus* (Bate, 1888) (AY853470.1) and *Litopenaeus vannamei* (Boone, 1931) (GQ857086.1) was 24.50%, whereas the average dissimilarity of COI gene of *P. monodon* (KF604891.1) with *F. indicus* (HM214712.1), *F. chinensis* (EU366250.1), *M. japonicus* (AY787755.1) and *L. vannamei* (AY781297.1) was 14.25%. The molecular clock for CR region was estimated as 3.44% per Myr from the ratio of the dissimilarities between CR and COI regions, or (24.5/14.25*2). The molecular clock rate for the 16S rRNA was calibrated in the same way and resulted in 1.12% per Myr, where the average dissimilarity of 16S rRNA gene between *P. monodon* (EU105473.1) and *F. indicus* (FJ002574.1), *F. chinensis* (AY264908.1), *M. japonicus* (EU056321.1) and *L. vannamei* (HQ127458.1) was 8.00%.

Origin of Bangladesh *Penaeus monodon* and phylogeography in the Indo-West Pacific region

The origin of Bangladesh *P. monodon* and the phylogeographic patterns were studied by constructing phylogenetic trees and by analysing genetic distances between the population samples. In total, 840 sequences for the CR of *P. monodon*, including 754 from GenBank and 86 from Bangladesh, were used to reconstruct the mtDNA phylogeny. The sequences were aligned using ClustalW Multiple alignment of Bioedit Sequence Alignment Editor (Hall, 1999) and were trimmed to 496 bp fragment, the sequences from the study of Walther et al. (2011) and Mkare et al. (2014) were aligned for 225 bps. The sequences of the CR were classified into the different lineages (A, B and C), as defined by Zhou et al. (2009) and Walther et al. (2011), by constructing a maximum

likelihood tree with partial deletion treatment for missing data and a general time reversible (GTR) model with gamma and invariant sites parameters (G+I) in MEGA ver. 6.06 (Tamura et al., 2013) (tree not shown). Divergence times of lineages A, B and C, based on the CR, were estimated in BEAST v1.7.5, using the calibrated molecular clock of 3.44% per Myr, and a constant population size as the tree prior. Markov chains were run for 1.0×10^7 generations and sampled every 1000th generation. Effective sample sizes (ESS) of the Markov Chain sampled model parameters were inspected using TRACER v1.5 (Rambaut & Drummond, 2009), and deemed sufficient if >200. Maximum clade credibility trees were constructed using TreeAnnotator v1.7.5 (Rambaut & Drummond, 2013) with a 10% burn-in. The divergence times were observed by reading Bayesian inference trees using the software FigTree v1.4.0 (Rambaut, 2012).

Phylogenetic analyses of the sequences from the CR lineage C were also performed separately as all Bangladesh sequences resolved onto this lineage. Trees were constructed with and without the shorter sequences from Indonesia, i.e. 225 bps instead of 496 bps. Three hundred and eleven sequences were obtained when omitting Indonesia and 337 sequences when including Indonesia. The phylogeny included 12 CR sequences from Kenya (You et al., 2008), and the shorter sequences from Mkare et al. (2014) were omitted. *Fenneropenaeus chinensis* was used as an outgroup for all analyses described above. Four different tree priors were evaluated for population growth, as implemented in BEAST: constant size with both strict clock and relaxed clock, exponential growth, logistic growth and expansion growth with strict clock.

The geographic patterns in variation of the CR were evaluated using 714 sequences from lineages A, B and C, and, separately, for 337 sequences from the C lineage (excluding the shorter sequences from the study of Mkare et al. 2014; see Table 1; Appendix 2). For both analyses, pairwise genetic distances (Φ_{ST}) based on the 496 bp fragment were calculated and tested by 1000 permutations using ARLEQUIN v3.5 (Excoffier & Lischer, 2011), except for comparisons with the Indonesian samples which were based on 225 bp. Multidimensional scale plots (Venables & Ripley, 2002) were drawn to visualize the differences among populations in R (R Core Team, 2014).

Similarly, pairwise genetic distances (Φ_{ST}) among 46 sequences of 16S rRNA from 6 populations (Bangladesh- 10, Sri Lanka- 5, India- 2, Malaysia- 4, China- 23 and Australia- 2 sequences) were calculated and tested by 1000 permutations using ARLEQUIN v3.5 (Excoffier & Lischer, 2011).

Orthologous or paralogous variation in *P. monodon* mitochondrial DNA?

To address the question of the paralogous duplications, raised by Walther et al. (2011), we searched for multiple peaks of the CR sequence electropherograms and by using two sets of primers (as described above). Furthermore, we evaluated whether the large split between the A and C lineages in the phylogeny of the CR region was observed also in phylogenies of the different mtDNA markers, the 16S rRNA gene and the combined fragment of the 16S rRNA and tRNA^{Val} genes, where we compared specimens sampled from the geographic regions which were characterized by the two lineages.

The 16S rRNA gene tree was constructed from 52 sequences (313 bps), containing 10 sequences from Bangladesh, five from Sri Lanka and 37 from other areas in the Indo-West Pacific (Table 1; Appendix 2). *Metapenaeus monoceros* (Fabricius 1798) (GenBank accession number: JX089983) was used as an outgroup. The combined 16S rRNA and tRNA^{Val} tree was constructed from 18 sequences (504 bps), including five sequences from Bangladesh, three from Sri Lanka and 10 available sequences from GenBank (Table 1 and Appendix 2). *Penaeus semisulcatus* (De Haan, 1844) (GenBank accession number: EF589704) was used as an outgroup. The trees were constructed following both strict clock and lognormal relaxed clock (uncorrelated) using BEAST as listed above for

the CR fragments and were redrawn in R (R Core Team, 2014) using the package APE (Paradis, 2006), and simplified by collapsing sequences from the same cluster and geographic area into a single taxonomic unit, labelled by the geographic location. Posterior probabilities (PP) were used to estimate the reliability of the internal nodes. ‘GTR’ was determined as the best substitution model for CR (C lineage), with parameters G for gamma-distributed rate variation among sites and I for invariant sites. The ‘TN93’ model with G was selected for the 16S rRNA and the combined fragment of the 16S rRNA and tRNA^{Val} genes. A calibrated molecular clock of 1.12% per Myr for the 16S rRNA gene and a combined fragment of the 16S rRNA and tRNA^{Val} region was used to calculate the divergence time. Variation of the different markers used for the phylogenetic trees, the CR, the 16S rRNA and the combined fragment of 16S rRNA and tRNA^{Val}, were summarized with nucleotide diversity (π) using ARLEQUIN v3.5 (Excoffier & Lischer, 2011). As for the CR trees, the 16S rRNA and combined 16S rRNA and tRNA^{Val} trees were also constructed using four priors of population growth.

Results

Mitochondrial DNA diversity of *P. monodon* within Bangladesh

Genetic diversity of the *P. monodon* CR sequences from Bangladesh is high, 124 segregating sites are observed and haplotype diversity reaches its maximum value ($h = 1.00$), as almost all haplotypes are singletons (83 haplotypes were found in 86 specimens) (Table 2). Four indels were observed, ranging in length of 1–2 bp, where two were polymorphic. One

Table 2 Genetic diversity of *P. monodon* from Bangladesh based on 86 sequences (520 bps) of mitochondrial CR

Sampling Location	<i>N</i>	<i>N_h</i>	<i>H_R</i>	<i>h</i> ± SE	π ± SE	<i>S</i>
SB Sundarban, Sathkhira	17	16	9.26	0.99 ± 0.02	0.027 ± 0.014	61
BC Barguna coast	10	10	10.00	1.00 ± 0.04	0.030 ± 0.016	49
MB Middle ground, Bay of Bengal	32	32	9.90	1.00 ± 0.01	0.022 ± 0.011	68
SM Sent Martin’s Island	27	27	10.00	1.00 ± 0.01	0.024 ± 0.012	74
All	86	83	10.00	1.00 ± 0.00	0.024 ± 0.012	124

N no. of individuals, *N_h* no. of haplotypes, *H_R* haplotype richness, *H* haplotype diversity, π nucleotide diversity, *S* no. of segregating sites, *SE* standard error

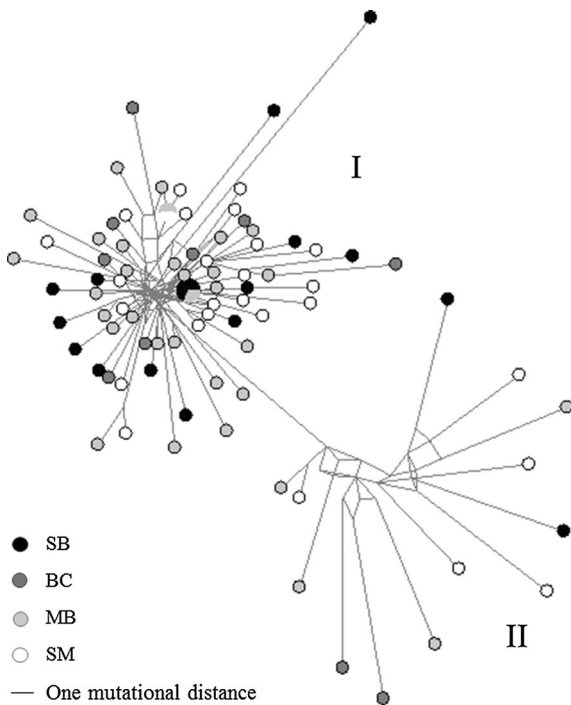


Fig. 2 Median-joining haplotype network based on CR of mtDNA of 86 individual *P. monodon* sampled from four locations along the Bangladesh coastline. Each pie represents a haplotype and its size reflects the frequency of samples. Distances between pies correspond to mutational differences between the haplotypes. Shading (from black to white) denotes different sampling locations (*SB*, *BC*, *MB* and *SM*; see Fig. 1), respectively

haplotype was identified in three specimens from locations *SB* (2 ind.) and *MB*, and another haplotype in two specimens from locations *MB* and *SM* (Fig. 2).

The CR haplotype network (Fig. 2) reflects the high variation, with two distinct clusters (I and II) separated by 13 mutations but without any geographic structure among the four sampling locations along the Bangladesh coastline. No differentiation was observed among the Bangladesh samples with Analysis of Molecular Variance (AMOVA) ($\Phi_{ST} = -0.0045$, $P = 0.60$), and comparison of the frequencies of the two clusters among the four sampling sites did not reveal any significant structure ($P = 0.86$). The haplotype network has several alternative connections between haplotypes at the centre of the two clusters. This reflects the large number of haplotypes segregating within the population with several homoplasious mutations. Forty nine of the variable nucleotide sites have more than two different nucleotides.

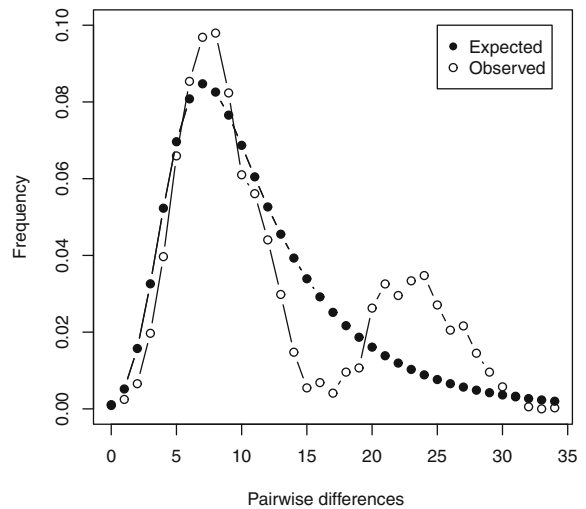


Fig. 3 Mismatch Distributions, based on 86 sequences (517 bps) of mitochondrial CR of *P. monodon* from Bangladesh, under the Sudden Expansion Model

The high haplotype diversity in Bangladesh is similar as observed elsewhere in the Indo-West Pacific region where it ranges from 0.97 to 1.00 (You et al., 2008; Zhou et al., 2009; Waqairatu et al., 2012; Mkare et al., 2014). The nucleotide diversity (π) in Bangladesh is also comparable to the most samples with a single CR lineage, and it is close to the average of the samples characterized with the C lineage but nominally lower than within the samples characterized by the A lineage, south and east of the Sunda Shelf (Table 3). Variation in samples where the distinct lineages A, B and C are found were considerably larger than in Bangladesh (Table 3).

The mismatch distribution for the Bangladesh *P. monodon* samples (Fig. 3) followed the sudden expansion model both for the SSD and the raggedness index ($P > 0.40$), despite clear observed bimodality, with a time of expansion at 143 (95% CI 90–397) Kyr. Deviation from the equilibrium was also observed with the Tajima's *D* and Fu's *F_s*, which were both negative and significant (Tajima's *D* = -1.72 , $P = 0.011$; Fu's *F_s* = -24.29 , $P < 0.001$), suggesting an expansion from a bottleneck or a selective sweep. The BSP analysis suggests that the Bangladesh population has a current effective population size (N_e) of 102.8×10^6 (CI 22.7×10^6 – 411.5×10^6) individuals and has undergone a gradual increase in N_e for the last ~ 200 – 350 Kyr (Fig. 4).

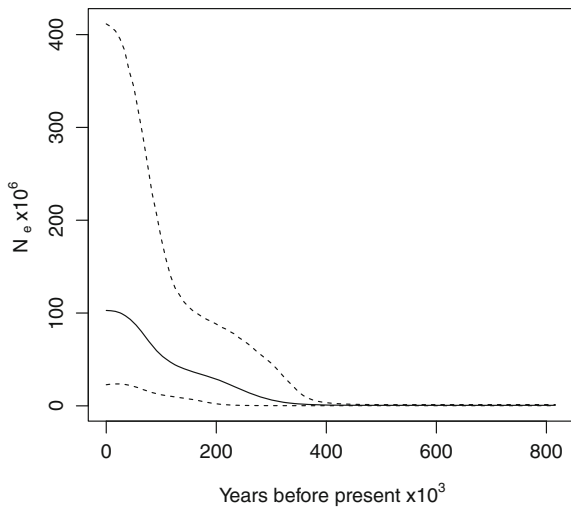


Fig. 4 Bayesian skyline plot showing the past population dynamics of *P. monodon* in Bangladesh estimated from 86 (517 bps) sequences of CR. Dotted lines represent the 95% confidence intervals. Effective population size ($N_e \times 10^6$) per generation is traced back in time from the present to the past

Origin and phylogeography

The most recent common ancestor of the three CR lineages dates back to 6.6 (CI 4.8–8.7) Mya, and 4.1 (CI 2.8–5.4) Mya for lineages A and B (tree not shown). The distribution of the CR lineages showed clear patterns in the Indo-West Pacific region longitudinally (Table 3). Sequences in lineage C were found in all specimens from Madagascar to Bangladesh, except for two specimen from Sri Lanka, and decreased in frequency in south-east Asia, from 50% in Thailand to 12% in China (Table 3). Lineages A and B were solely found in populations from the south-east Asia and the Pacific Ocean regions (Table 3).

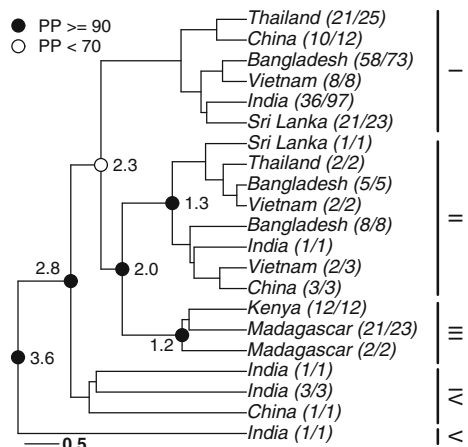


Fig. 5 Bayesian inference tree based on 311 CR (496 bps) sequences of *P. monodon* of lineage C (without Indonesia) under ‘GTR+G+I’ model following strict clock of 3.44% per Myr. Numbers at the nodes represent divergence in million years and shadings represent Bayesian posterior probabilities (PP, %). The tree is rooted with *F. chinensis* (GenBank accession number: DQ518969) with a divergence time of ~46.2 (CI 32.3–65.0) Mya (PP ≥ 90). Numbers following names of countries represent the number of haplotypes/sample size. Note: the clustering within clades I to III is insignificant except for a single node with PP ≥ 90, and different haplotypes from the same country have been grouped into a single branch

The Bayesian inference analyses of lineage C of the CR, showed a complex pattern in the Indo-West Pacific region. The tree had five clusters (PP ≥ 90, Fig. 5). Haplotypes from Bangladesh are characterized by the same notations (I and II) as in Fig. 2. A tree based on shorter sequences that included Indonesia produced a less-supported tree where the Indonesian sequences aggregate within clusters I and II (tree not shown). The earliest splits [~3.6 (CI 2.6–4.3) Mya,

Table 3 Population and lineage-wise distribution of CR sequences of *P. monodon* from the Indo-West Pacific region

CR lineage	MG	KY	IW	IS	IE	SL	BD	TW	TE	ID	VN	CH	PH	AU	Total
A						2		17	15	111	47	113	35	29	369
B									1	5			1		7
C	25	138	2	23	83	24	86	11	16	26	13	16			463
Other										1					1
All	25	138	2	23	83	26	86	28	32	143	60	129	36	29	840
% C	100	100	100	100	100	92	100	39	50	18	22	12	0	0	
π	0.026	0.020	0.015	0.022	0.051	0.038	0.026	0.086	0.089	0.046	0.070	0.054	0.036	0.034	

The nucleotide diversity (π) is based on 496 bps, except for KY (379 bps) and ID (226 bps). Information of the samples is given in Table 1 and Appendix 2. Letters: MG Madagascar, KY Kenya, IW India West, IS India South, IE India East, SL Sri Lanka, BD Bangladesh, TW Thailand West, TE Thailand East, ID Indonesia, VN Vietnam, CH China, PH Philippines, AU Australia

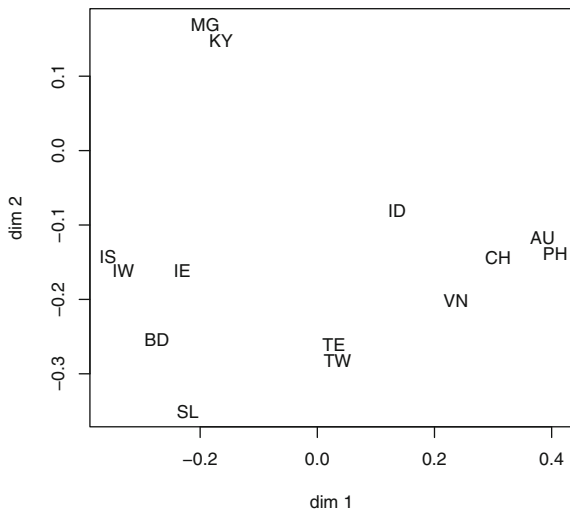


Fig. 6 Multidimensional scale plot based on pairwise genetic distances (Φ_{ST}) between 14 *P. monodon* populations, including variation in 714 sequences of all mtDNA CR lineages in the Indo-West Pacific region. Letters: MG Madagascar, KY Kenya, IW India West, IS India South, IE India East, SL Sri Lanka, BD Bangladesh, TW Thailand West, TE Thailand East, ID Indonesia, VN Vietnam, CH China, PH Philippines, AU Australia

cluster IV–V] within lineage C were found for unpublished CR sequences from the east coast of India and one sequence from China ($PP \geq 90$, Fig. 5).

The CR trees (lineage C) obtained with the different priors showed similar topologies with slight variation in divergence time, except for the tree based on relaxed clock which had wider confidence intervals (Appendix 3).

Nucleotide diversity of all CR sequences was 0.091, substantially larger than within the C lineage where it was 0.040, reflecting the deep divergence of the A, B and C lineages. When excluding the highly divergent lineage from Kenya and Madagascar the diversity within the C lineage lowered to 0.032, but it was still higher than within Bangladesh (0.024) pointing to population structure in the Indian Ocean, as described further below. Less nucleotide diversity was found within the 16S rRNA gene (0.024) and the combined fragment of 16S rRNA and tRNA^{Val} (0.035) genes than for the CR sequences. Even though the samples do not cover the same geographic range, the two datasets include samples from the geographic areas characterized by the distinct A (Australia) and C (e.g. Bangladesh) CR lineages.

Pairwise genetic distances (Φ_{ST}) for CR showed clear differentiation between populations of all lineages (Fig. 6; Table 4) and also within lineage C. The Madagascar and Kenya populations, which cluster within the C lineage, are closely related to each other ($\Phi_{ST} = 0$) but significantly different from other

Table 4 Pairwise genetic distances (Φ_{ST}) between *P. monodon* populations in the Indo-West Pacific region

	MG	KY	IW	IS	IE	SL	BD	TW	TE	ID	VN	CH	PH	AU
MG	*	0.00	0.62	0.65	0.48	0.56	0.62	0.55	0.49	0.72	0.63	0.69	0.81	0.82
KY	0.31	*	0.65	0.67	0.47	0.54	0.62	0.50	0.44	0.70	0.60	0.68	0.80	0.81
IW	0.00	0.01	*	-0.1	-0.20	-0.18	-0.03	0.27	0.13	0.60	0.49	0.61	0.77	0.78
IS	0.00	0.00	0.68	*	-0.01	0.03	0.11	0.45	0.33	0.65	0.58	0.65	0.80	0.81
IE	0.00	0.00	0.76	0.82	*	0.02	0.08	0.43	0.32	0.67	0.56	0.63	0.71	0.71
SL	0.00	0.00	0.84	0.04	0.06	*	0.09	0.36	0.24	0.60	0.51	0.60	0.74	0.74
BD	0.00	0.00	0.44	0.00	0.00	0.00	*	0.53	0.41	0.70	0.63	0.68	0.80	0.81
TW	0.00	0.00	0.04	0.00	0.00	0.00	0.00	*	-0.01	0.09	0.03	0.11	0.21	0.26
TE	0.00	0.00	0.37	0.00	0.00	0.00	0.00	0.36	*	0.20	0.11	0.22	0.32	0.35
ID	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	*	0.02	0.02	0.07	0.13
VN	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.11	0.00	0.10	*	0.01	0.07	0.15
CH	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.14	*	0.02	0.13
PH	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.05	*	0.11
AU	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	*

The distances, based on 714 sequences, are shown above the diagonal. The respective *P* values, obtained by 1000 permutations, are below the diagonal (see Materials and methods for further details). Letters refer to populations as listed in Table 3. As the IW sample is only based on two individuals the estimates including comparison with IW (in italic) are unreliable

populations (Φ_{ST} ranges from 0.44 to 0.82, $P < 0.01$; Table 4). The separation among the other populations follow the first discriminant axis from India in the west, to the Philippines in the east (together with the Australian sample) (Fig. 6). Differences among populations including lineages A, B and C are almost all significant except where samples harbour large variation resulting in lower Φ_{ST} between populations and reduced statistical power, as between Thailand (W) and Thailand (E), between Vietnam and Indonesia and between Vietnam and China. The Bangladesh population is most similar to the populations of India and Sri Lanka ($\Phi_{ST} = 0.08$ –0.11, omitting the small sample from IW) but significantly different from east and south India (Fig. 6; Table 4). The differentiation of the Bangladesh population from Sri Lanka, Thailand (W and E), Indonesia and China were insignificant when lineages A and B are omitted (Φ_{ST} ranges from 0 to 0.03, $P > 0.05$). The distances between populations based on the 16S rRNA gene support the split based on CR between the Bangladesh population and the populations from the Pacific region (Malaysia, China and Australia, $P < 0.05$) and did not reveal differentiation from India and Sri Lanka ($P > 0.05$).

Orthologous or paralogous variation

No indications were observed of different sequences within individuals from Bangladesh, which could suggest heteroplasmy or duplications, even though the same individuals were amplified for the short and the long fragment of the CR region, as reported for the Indonesian population by Walther et al. (2011). The phylogenetic analyses of the two other mtDNA markers supported the overall split observed among the A, B and C lineages of the CR, although the time estimates differed (Figs. 7, 8). The 16S rRNA tree has three clusters ($PP \geq 90$) for both the strict clock and the lognormal relaxed clock models; cluster A with sequences from the western Pacific region deviated from clusters C1 and C2, with sequences from all locations of Indo-West Pacific region (Fig. 7) for about 2.6 (CI 1.5–3.6) Mya. None of the samples represent localities characterized by the CR lineage B in this analysis. Similarly, the combined fragment of 16S rRNA and tRNA^{Val} genes produced three distinct clusters for both strict clock and lognormal (uncorrelated) relaxed clock ($PP \geq 90$) (Fig. 8). Cluster A and B were composed of samples from the western Pacific

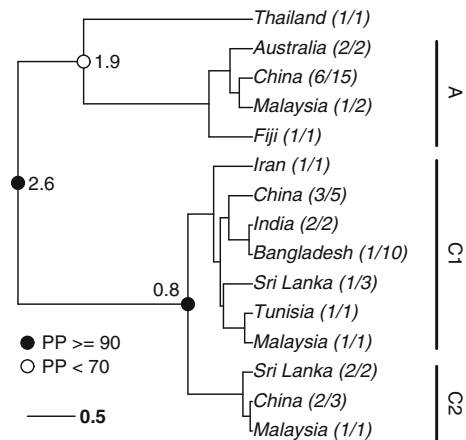


Fig. 7 Bayesian inference tree based on 16S rRNA gene of *P. monodon* under ‘TN93 + G’ model following strict clock of 1.12% per Myr. Numbers at the nodes represent divergence in million years and shadings represent Bayesian posterior probabilities (PP, %). The tree is rooted with *M. monoceros* (GenBank accession number: JX089983) with a divergence time of ~14.0 (CI 9.9–18.8) Mya ($PP \geq 90$). Numbers following names of countries represent the number of haplotypes/sample size

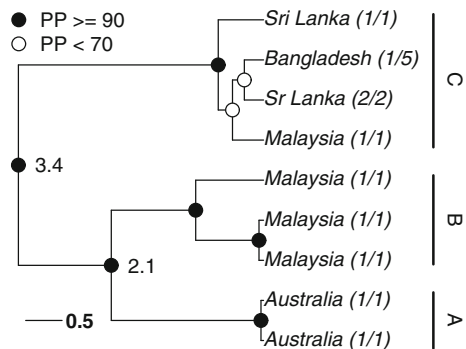


Fig. 8 Bayesian inference tree based on a combined fragment of 16S rRNA and tRNA^{Val} genes of *P. monodon* under ‘TN93 + G’ model following strict clock of 1.12%. Numbers at the nodes represent divergence in million years and shadings represent Bayesian posterior probabilities (PP, %). The tree is rooted with *P. semisulcatus* (GenBank accession number: EF589704) with a divergence time of ~14.3 (CI 10.0–19.8) Mya ($PP \geq 90$). Numbers following names of countries represent the number of haplotypes/sample size

region (Australia) and south-east Asia (Malaysia), respectively. Cluster C comprised samples from Bangladesh and Sri Lanka and also a specimen from Malaysia. Topology and the age of nodes of the trees constructed with the different priors for population

growth were similar, except for the tree based on relaxed clock which had wider confidence intervals for the CR C lineage, and exceeded the wider interval obtained when including the lower and upper rate for the strict mutation clock for the CR and the constant size model (Appendix 3).

Discussion

Mitochondrial DNA in *P. monodon* from Bangladesh is characterized by a high degree of variation where almost all individuals carry a unique haplotype. High genetic diversities have also been reported in *P. monodon* populations from the Indo-West Pacific region (You et al., 2008; Zhou et al., 2009; Waqairatu et al., 2012; Mkare et al., 2014). The distribution of the mtDNA variation along the Bangladesh coast does not provide any evidence of population structure, reflecting a random mixing of pelagic larval stages and or migration of adults. A single population in Bangladesh based on mtDNA variation was also observed in *F. indicus* (Alam et al., 2015). The high mitochondrial variation in *P. monodon* in Bangladesh and the two distinct clusters indicates that the population size is large and may have a long history (Rogers & Harpending, 1992), but the mismatch analyses support a sudden expansion around 90–397 Kyr ago. A similar result was obtained with the Bayesian skyline plot method, indicating growth over the last 200–350 Kyr, well before the onset of last glacial epoch of the Pleistocene and earlier than the expansion in *F. indicus* from Bangladesh (Alam et al., 2015). This time estimate may though be an overestimate due to a possible admixture of distinct mtDNA clusters, as in Grant et al. (2012) who demonstrated also how decline in population size may eradicate historical information; thus the interpretation of estimates of demographic changes should be taken with caution. Large variation was though observed within both clusters, where distinct haplotypes were found and the high number of homoplasies may further underestimate the number of mutations that have occurred. More markers are needed to evaluate this time interval, suggested by the analyses of the mtDNA and to obtain more precise estimates.

The overall genetic patterns based on all CR sequences closely follows the geographic structure in the Indo-West Pacific region, where neighbouring

populations are generally more similar to each other than those sampled further away. The Bangladesh *P. monodon* population shows the greatest similarities with the neighbouring populations in India and Sri Lanka but differs clearly from all other samples. Such a pattern has also been observed in a *Macrobrachium rosenbergii* (De Man 1879) population from Bangladesh, which showed closest similarity with the Indian and Sri Lankan population (Hurwood et al., 2014). Madagascar and Kenya populations, coming from a different biogeographical province, the Western Indian Province (Bowen et al., 2012), were distinct from all other populations. The south-east Asian populations differed from the south Asian populations and the south-east African populations, but were somewhat related to the West Pacific populations. Biogeographic events related to the Sunda Shelf are known to have affected diversification of marine organisms in the Indo-West Pacific region (Gopurenko et al., 1999; Tsoi et al., 2007), and have been proposed to explain the splits (among South Pacific, eastern Australia and south-east Asia) and a split (between eastern Indian Ocean and western Pacific) within *P. monodon* as the two oceans were completely separated at certain times during the Tertiary and Pleistocene periods (Waqairatu et al., 2012). The oldest divergence within the species (estimated as the most recent common ancestor for all lineages), based on the mutation rate for the CR, predates the main temperature fluctuation and sea level changes which started around 2.8 Mya or at the onset of the Pleistocene, when the Indian and Pacific Oceans were completely separated (Voris, 2000; Benzie et al., 2002). The split between lineage C, which characterizes the region west of the Sunda Shelf, and lineage A, the region east of the Sunda Shelf, dates back to 6.6 (CI 4.8–8.7) Mya and lineages B and A diverged from each other before or at the onset of the Pleistocene, 4.1 (CI 2.8–5.4) Mya. The estimates are though dependent on the mutation rate and are older than the estimates obtained for the 16S rRNA and the tRNA genes. Concerns have been raised about using estimates of molecular clock rates from comparisons between species to make inferences about divergence of populations within species as they can lead to overestimation of the times (Ho et al., 2005, 2007), as not all of the variation may become fixed between species. However, the time estimates observed in this study are generally large and exceed the period of 1 Myr where this may be of

main concern (Papadopoulou et al., 2010). An admixture of all lineages may have occurred around the Sunda Shelf in Indonesia, Thailand, Vietnam and China, when the Indian and Pacific Oceans connected through the Sunda Shelf or even after the last glacial period (Bird et al., 2005). The admixture appears to have been directional, as lineage C from the Bay of Bengal area, rather than from the distinct African populations as proposed by Benzie et al. (2002), may have migrated eastwards through the region, but only a small number of lineage A individuals have been found west of the region, in western Thailand and Sri Lanka. Baldwin et al. (1998) concluded that the *Penaeus* genus originated in the centre of the Indo-West Pacific region and migrated eastwards and westwards forming two groups, in the eastern Pacific and the Indian Ocean. Similar diversification may have occurred within *P. monodon*, but despite these larger overall patterns, further splits are observed within lineage C, which comprises three main clusters: one with south-east Africa populations and two others with mixed populations from south and south-east Asia and from the Pacific Ocean region. The mixed populations, forming clusters that diverged approximately at 2.3 (CI 1.7–2.6) Mya, could possibly reflect isolation due to sea level changes during the Pleistocene in the Sunda Shelf. Such divergence has been reported for other Indo-West Pacific species (see Alam et al., 2015), e.g. in mud crab (*Scylla serrata* Forskal, 1775) (Gopurenko et al., 1999), kuruma shrimp (*Penaeus japonicus* Bates 1888) (Tsoi et al., 2007). The Sunda Shelf may not be the only geographic barrier in the Indian Ocean. The large split between the African samples and the samples from India, Sri Lanka and Bangladesh might have occurred due to large geographical distances, and due to different surface and subsurface equatorial ocean currents in western Indian Ocean (Pidwirny, 2006). Similarly, three distinct evolutionary lineages among *F. indicus* populations were recently reported west of the Sunda Shelf region: one in the Western Indian Ocean and Thailand and two in the eastern Indian Ocean, off Bangladesh and Sri Lanka, which might result from vicariance both before and during Pleistocene (Alam et al., 2015).

Walther et al. (2011) detected different sequences within a subset of individuals of *P. monodon* from Indonesia and concluded that they could have arisen due to duplication, where lineage C comprised

paralogs, either due to insertion in the nuclear genome (numts) or within the mtDNA. The authors argued that previous studies had thus highly overestimated the variation within the mtDNA. No indications of numts, duplications or heteroplasmy were found in our sequences, as evidenced by multiple peaks in the electropherograms or different sequences from the same individuals. In addition, genealogies based on the other two mtDNA markers, 16S rRNA, and a fragment of 16S rRNA and tRNA^{Val} genes support the split between the C and the A lineages, as seen in comparisons of specimens from Australia (CR lineage A) and specimens from Bangladesh (CR lineage C). Although the estimated divergence time differed for the different trees, the confidence intervals were wide and the exact mutation rate of the different markers may also be uncertain. A split between populations from the eastern Indian Ocean and western Pacific has also been observed for the elongation factor 1-alpha gene (Duda & Palumbi, 1999). Sequences of CR from Kenya (Mkare et al., 2014), which clustered in lineage C, from 129 individuals (126 haplotypes) of *P. monodon*, did not either show any evidence of co-amplification of pseudogenes and/or paralogous genes. Walther et al. (2011) acknowledged that the occurrence of such divergent lineages could be due to heteroplasmy as a result of bi-parental inheritance. Heteroplasmy has been described for several species such as Atlantic salmon (*Salmo salar* Linnaeus, 1758) and brown trout (*Salmo trutta* Linnaeus, 1758) (Ciborowski et al., 2007) and among *Drosophila* species (Wolff et al., 2013; Dokianakis & Ladoukakis, 2014), and may be more likely to be detected in hybridizing species due to higher sequence divergence or higher paternal leakage. The distinct lineages A, C are likely to characterize previously isolated groups within *P. monodon*, or even cryptic species as in *F. indicus* (Alam et al., 2015). A secondary contact zone of the two lineages may exist in Indonesia and South-East Asia where paternal leakage may occur, resulting in heteroplasmy within individuals, as observed by Walther et al. (2011).

To conclude, this study has unveiled that the Bangladesh *P. monodon* population is large and has high genetic diversity, which is in accordance with populations of this species throughout its distribution. The Bangladesh population, although genetically differentiated, was most similar to the populations in India and Sri Lanka which belong to lineage C, one of

three main mtDNA lineages within the species. The South-East Asia region was identified as containing an admixture of all lineages, whereas the western Indian Ocean and western Pacific Ocean regions were dominated by lineages C and A, respectively, but distinct population structure also occurred within the Western Indian Ocean. High genetic diversity in Bangladesh *P. monodon* and different lineages in the Indo-West Pacific region should be considered for proper management of genetic diversity and for aquaculture development of the species. This may be especially of concern where only one of the lineages is found, such as in the Bay of Bengal, where mixing of the different lineages should be avoided. The distinct lineages within the species may even represent cryptic species, considering the large divergence times that predate the split of many other species (e.g. Alam et al., 2015). Further analyses, based on nuclear genomic markers, are warranted to evaluate this evolutionary split of the mtDNA lineages; whether it presents cryptic species or a paralogous event. Breeding experiments among the lineages could provide further evidence of their divergence and possible paternal mtDNA transmission. The high genetic diversity and geographic patterns suggest that *P. monodon* is not threatened, but several concerns due to harvesting pressure, aquaculture and habitat destruction need to be considered.

Acknowledgments The study was funded by the United Nations University–Fisheries Training Programme (UNU-FTP), Iceland. The first author was supported under a doctoral scholarship by the UNU-FTP, Iceland. The authors are grateful to UNU-FTP, Iceland, for its support and the Department of Fisheries, Bangladesh, for its assistance during shrimp sample collection. We thank the editors and two anonymous reviewers for valuable comments and suggestions.

References

- Alam, M. M. M., K. M. Westfall & S. Pálsson, 2015. Mitochondrial DNA variation reveals cryptic species in *Fenneropenaeus indicus*. *Bulletin of Marine Science* 91(1): 15–31.
- Baldwin, J. D., A. L. Bass, B. W. Bowen & W. H. Clark, 1998. Molecular phylogeny and biogeography of the marine shrimp *Penaeus*. *Molecular Phylogenetics and Evolution* 10: 399–407.
- Bandelt, H. J., P. Forester & A. Röhl, 1999. Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16: 37–48.
- Benzie, J. A. H., E. Ballment, A. T. Forbes, N. T. Demetriades, K. Sugama, Haryanti & S. Moria, 2002. Mitochondrial DNA variation in Indo-Pacific populations of the giant tiger prawn, *Penaeus monodon*. *Molecular Ecology* 11(12): 2553–2569.
- Benzie, J. A. H., E. Ballment & S. Frusher, 1993. Genetic structure of *Penaeus monodon* in Australia: concordant results from mtDNA and allozymes. *Aquaculture* 111(1–4): 89–93.
- Benzie, J. A. H., S. Frusher & E. Ballment, 1992. Geographical variation in allozyme frequencies of populations of *Penaeus monodon* (Crustacea: decapoda) in Australia. *Australian Journal of Marine and Freshwater Research* 43(4): 715–725.
- Bird, M. I., D. Taylor & C. Hunt, 2005. Palaeoenvironments of insular Southeast Asia during the Last Glacial Period: a savanna corridor in Sundaland? *Quaternary Science Reviews* 24(2005): 2228–2242.
- Bowen, B. W., L. A. Rocha, R. J. Toonen, S. A. Karl & the ToBo Laboratory, 2012. The origins of tropical marine biodiversity. *Trends in Ecology & Evolution* 28(6): 359–366.
- Chu, K. H., C. P. Li, Y. K. Tam & S. Lavery, 2003. Application of mitochondrial D-loop region in population genetic studies of the shrimp *Penaeus*. *Molecular Ecology Notes* 3: 120–122.
- Ciborowski, K. L., S. Consuegra, C. G. de Leániz, M. A. Beaumont, J. Wang & W. C. Jordan, 2007. Rare and fleeting: an example of interspecific recombination in animal mitochondrial DNA. *Biology Letters* 3: 554–557.
- Department of Fisheries Bangladesh, 2013. National Fish Week 2013 Compendium (in Bengali). Department of Fisheries, Ministry of Fisheries and Livestock, Bangladesh.
- Dokianakis, E. & E. D. Ladoukakis, 2014. Different degree of paternal mtDNA leakage between male and female progeny in interspecific *Drosophila* crosses. *Ecology and Evolution* 4(13): 2633–2641.
- Drummond, A. J., S. Y. W. Ho, N. Rawlence & A. Rambaut, 2007. A Rough Guide to BEAST 1.4. <http://beast.bio.ed.ac.uk>.
- Duda Jr, T. F. & S. R. Palumbi, 1999. Population structure of the black tiger prawn, *Penaeus monodon*, among western Indian Ocean and western Pacific populations. *Marine Biology* 134: 705–710.
- Excoffier, L. & H. Lischer, 2011. Arlequin ver 3.5: An Integrated Software Package for Population Genetics Data Analysis. Swiss Institute of Bioinformatics, Switzerland.
- Forbes, A. T., N. T. Demetriades, J. A. H. Benzie & E. Ballment, 1999. Allozyme frequencies indicate little geographic variation among stocks of giant tiger prawn *Penaeus monodon* in the south-west Indian Ocean. *South African Journal of Marine Science* 21(1): 271–277.
- Fu, Y. X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147: 915–925.
- Gaither, M. R., B. W. Bowen, T. Bordenave, L. A. Rocha, S. J. Newman, J. A. Gomez, L. van Herwerden & M. T. Craig, 2001. Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the indo-pacific barrier with contemporary overlap in the coral triangle. *BMC Evolutionary Biology* 2011(11): 189.

- Gopurenko, D., J. M. Hughes & C. P. Keenan, 1999. Mitochondrial DNA evidence for rapid colonisation of the Indo-West Pacific by the mud crab *Scylla serrata*. *Marine Biology* 134: 227–233.
- Goudet, J., 2005. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes* 5(1): 184–186.
- Grant, W. S., M. Liu, T. X. Gao & T. Yanagimoto, 2012. Limits of Bayesian skyline plot analysis of mtDNA sequences to infer historical demographics in Pacific herring (and other species). *Molecular Phylogenetics and Evolution* 65(2012): 203–212.
- Guindon, S., J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk & O. Gascuel, 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: assessing the Performance of PhyML 3.0. *Systematic Biology* 59(3): 307–321.
- Hall, T. A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41: 95–98.
- Harpending, R. C., 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology* 66: 591–600.
- Ho, S. Y. W., M. J. Phillips, A. Cooper & A. J. Drummond, 2005. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution* 22(7): 1561–1568.
- Ho, S. Y. W., B. Shapiro, M. J. Phillips, A. Cooper & A. J. Drummond, 2007. Evidence for time dependency of molecular rate estimates. *Systematics Biology* 56: 515–522.
- Hurwood, D. A., S. Dammanagoda, M. N. Krosch, H. Jung, K. R. Salin, M. A.-B. H. Youssef, M. de Bruyn & P. B. Mather, 2014. Impacts of climatic factors on evolution of molecular diversity and the natural distribution of wild stocks of the giant freshwater prawn (*Macrobrachium rosenbergii*). *Freshwater Science* 33(1): 217–231.
- Keyse, J., E. D. Crandall, R. J. Toonen, C. P. Meyer, E. A. Trembl & C. Riginos, 2014. The scope of published population genetic data for Indo-Pacific marine fauna and future research opportunities in the region. *Bulletin of Marine Science* 90: 47–78.
- Knowlton, N. & L. A. Weigt, 1998. New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society London B* 265: 2257–2263.
- Knowlton, N., L. A. Weigt, L. A. Solorzano, D. K. Mills & E. Bermingham, 1993. Divergence in proteins, mitochondrial DNA, and reproductive compatibility across the Isthmus of Panama. *Science* 260: 1629–1632.
- Kumar, N., W. S. Lakra, K. C. Majumdar, M. Goswami & K. Ravinder, 2007. Genetic diversity in the Indian population of *Penaeus monodon* (Fabricius, 1798) as revealed by mtDNA sequence analysis. *Aquatic Resources* 38: 862–869.
- Lavery, S., T. Y. Chan, Y. K. Tam & K. H. Chuc, 2004. Phylogenetic relationships and evolutionary history of the shrimp genus *Penaeus s.l.* derived from mitochondrial DNA. *Molecular Phylogenetics and Evolution* 31: 39–49.
- MacDonald, K. S., L. Yampolsky & J. E. Duffy, 2005. Molecular and morphological evolution of the amphipod radiation of Lake Baikal. *Molecular Phylogenetics and Evolution* 35: 323–343.
- Mandal, A., D. Rao, D. Karuppaiah, A. Gopalakrishnan, J. Pozhoth, Y. Chellamma, T. Samraj & R. W. Doyle, 2012. Population genetic structure of *Penaeus monodon*, in relation to monsoon current patterns in Southwest, East and Andaman coastal waters of India. *Gene* 491: 149–157.
- Michel-Salzat, A. & D. Bouchon, 2000. Phylogenetic analysis of mitochondrial LSU rRNA in oniscids. *CR De L Académie Des Science Series III-Science De La Vietnam-life Science* 323(9): 827–837.
- Mkare, T. K., S. von der Heyden, J. C. Groeneveld & C. A. Matthee, 2014. Genetic population structure and recruitment patterns of three sympatric shallow-water penaeid prawns in Ungwana Bay, Kenya, with implication for fisheries management. *Marine and Freshwater Research* 65(3): 255–266.
- Motoh, H., 1985. Biology and ecology of *Penaeus monodon*, in: Taki, Y., J. H. Primavera, & Llobrera J.A. (eds), *Proceedings of the First International Conference on the Culture of Penaeid Prawns/Shrimps*. Iloilo City.
- Oanh, K. T. P., T. T. Thanh, N. D. Ton, H. T. T. Hue, N. H. Ninh, P. A. Tuan & N. V. Hai, 2010. Mitochondrial haplotype diversity of black tiger shrimp (*Penaeus monodon*) in Vietnam. *Tap Chi Cong Nghe Sinh Hoc* 8(4): 1801–1808.
- Palumbi, S., A. Martin & S. Romano, 1991. *The Simple Fool's Guide to PCR*. Department of Zoology and Kewalo Marine Laboratory. University of Hawaii, Honolulu.
- Papadopoulou, A., I. Anastasiou & A. P. Vogler, 2010. Revisiting the insect mitochondrial molecular clock: the mid-aegean trench calibration. *Molecular Biology and Evolution* 27(7): 1659–1672.
- Paradis, E., 2006. *Analysis of Phylogenetics and Evolution with R*. © 2006 Springer Sci Bus Media, LLC.
- Pascoal, A., I. Ortea, J. M. Gallardo, B. Canas, J. Barros-Velazquez & P. Calo-Mata, 2012. Species identification of the Northern shrimp (*Pandalus borealis*) by polymerase chain reaction-restriction fragment length polymorphism and proteomic analysis. *Analytical Biochemistry* 421(1): 56–67.
- Pascoal, A., J. Barros-Velazquez, A. Cepeda, J. M. Gallardo & P. Calo-Mata, 2008. A polymerase chain reaction-restriction fragment length polymorphism method based on the analysis of a 16S rRNA/tRNA(Val) mitochondrial region for species identification of commercial penaeid shrimps (Crustacea: Decapoda: Penaeoidea) of food interest. *Electrophoresis* 29: 499–509.
- Pidwirny, M. 2006. Surface and Subsurface Ocean Currents: Ocean Current Map. *Fundamentals of Physical Geography*, 2nd Edition. Date: 14.04.2015. http://www.physicalgeography.net/fundamentals/8q_1.html.
- Quader, O., 2010. Coastal and marine biodiversity of Bangladesh (Bay of Bengal). *Proc. Of International Conference on Environmental Aspects of Bangladesh (ICEAB10)*, Japan.
- R Core Team, 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rambaut, A., 2012. *FigTree v1.4* 2006–2012. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh.
- Rambaut, A. & A. J. Drummond, 2009. *MCMC Trace Analysis Tool*, Version v1.5.0, 2003–2009. Institute of Evolutionary

- Biology, University of Edinburgh. <http://code.google.com/p/beast-mcmc/>.
- Rambaut, A. & A. J. Drummond, 2013. TreeAnnotator v1.7.5, 2002–2013. Institute of Evolutionary Biology, University of Edinburgh and Department of Computer Science, University of Auckland. <http://beast.bio.ed.ac.uk/>.
- Rogers, A. R. & H. Harpending, 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9: 552–569.
- Tajima, F., 1989a. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585–595.
- Tajima, F., 1989b. The effect of change in population size on DNA polymorphism. *Genetics* 123: 597–601.
- Tajima, F., 1993. Measurement of DNA polymorphism. In Takahata, N. & A. G. Clark (eds), *Mechanisms of Molecular Evolution. Introduction to Molecular Paleopopulation Biology*. Japan Scientific Societies Press, Sinauer Associates, Inc., Sunderland: 37–59.
- Tamura, K., G. Stecher, D. Peterson, A. Filipski & S. Kumar, 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* 30: 2725–2729.
- Tong, J. G., T. Y. Chan & K. H. Chu, 2000. PA preliminary phylogenetic analysis of *Metapenaeopsis* (Decapoda: Penaeidae) based on mitochondrial DNA sequences of selected species from the Indo-West Pacific. *Crustacean Biology* 20(3): 541–549.
- Tsoi, K. H., T. Y. Chan & K. H. Chu, 2007. Molecular population structure of the kuruma shrimp *Penaeus japonicus* species complex in western Pacific. *Marine Biology* 150: 1345–1364.
- Vaseeharan, B., P. Rajakamaran, D. Jayaseelan & A. Y. Vincent, 2013. Molecular markers and their application in genetic diversity of penaeid shrimp. *Aquaculture International* 21: 219–241.
- Venables, W. N. & B. D. Ripley, 2002. *Modern Applied Statistics with S*. 4th edn, Springer, New York. ISBN 0-387-95457-0.
- Voris, H. K., 2000. Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. *Journal of Biogeography* 27: 153–167.
- Walther, E., G. Schöfl, G. Mrotzek, K. Sugama Haryanti & H. P. Saluz, 2011. Paralogous mitochondrial control region in the giant tiger shrimp, *Penaeus monodon* (F.) affects population genetics inference: a cautionary tale. *Molecular Phylogenetics and Evolution* 58: 404–408.
- Waqairatu, S. S., L. Dierens, J. A. Cowley, T. J. Dixon, K. N. Johnson, A. C. Barnes & Y. Li, 2012. Genetic analysis of Black Tiger shrimp (*Penaeus monodon*) across its natural distribution range reveals more recent colonization of Fiji and other South Pacific islands. *Ecology and Evolution* 2(8): 2057–2071.
- Wilson, K., V. Cahill, E. Ballment & J. Benzie, 2000. The complete sequence of the mitochondrial genome of the crustacean *Penaeus monodon*: are malacostracan crustaceans more closely related to insects than to branchiopods? *Molecular Biology and Evolution* 17(6): 863–874.
- Wolff, J. N., M. Nafisinia, P. Sutovsky & J. W. O. Ballard, 2013. Paternal transmission of mitochondrial DNA as an integral part of mitochondrial inheritance in metapopulations of *Drosophila simulans*. *Heredity* 110: 57–62.
- You, E. M., T. S. Chiu, K. F. Liu, A. Tassanakajon, S. Klinbunga, K. Triwitayakorn, L. D. de la Pena, Y. Li & H. T. Yu, 2008. Microsatellite and mitochondrial haplotype diversity reveals population differentiation in the tiger shrimp (*Penaeus monodon*) in the Indo-Pacific region. *Animal Genetics* 39: 267–277.
- Zhou, F. L., S. G. Jiang, Y. J. Jiang, J. H. Huang & Z. M. Ma, 2009. Population genetic structure of the tiger prawn (*Penaeus monodon*) in the coastal waters of South China, based on mitochondrial DNA control region sequences. *Applied Ichthyology* 25: 411–416.