Evidence of genetic homogeneity in longtail tuna in the west coast of India and the distantly set Andaman Archipelago.

Mohammed koya, K[1]., Sandhya Sukumaran[1], Prathibha Rohit[1], Abdussamad, E.M[1]. and Arun Ratheesh[2]

[1] ICAR-Central Marine Fisheries Research Institute, Kochi, India; [2] ICAR-Central Island Agricultural Research Institute, Port Blair, India

Introduction

Knowing the distribution of fish stocks is an important component for understanding the stock status to design and implement management regimes. There has been a wider debate on the stock structure of *Thunnus tonggol* (Bleeker, 1851) in the Indo-Pacific mainly based on the meristic and morphometric (Serventy, 1956; Gibbs & Collette, 1967; Silas, 1967; Wilson, 1981a and Abdulhaleem, 1989) as well as electrophoretic (Lewis, 1981) studies which postulated numerous stocks (self-sustaining units) throughout the distributional range of the species. Many of the studies (Hauser & Seeb 2008, Griffith *et al*., 2010b) have suggested use of molecular genetics techniques for unravelling stock structure in the individual basins as well as across considering its preciseness and reliability. Use of population genetics in fisheries research has increased dramatically in the past few decades owing to continuous emergence of new methods and decreasing costs (Kumar & Martin, 2015). Indian Ocean Tuna Commission (IOTC), the Regional Fisheries Management organization (RFMO) monitoring and managing the fisheries for tuna in the Indian Ocean urged (IOTC, 2017a) for concerted efforts from the CPCs for stock delineation using molecular techniques. Only recently, the studies have hinted on the stock structure of the species in the Indian Ocean (Kunal *et al.,* 2014) and South China Sea (Willette *et al.,* 2016) using genetic studies.

Management of LOT is of international importance owing to the species being a straddling resource across many developing nations. Being a neritic resource, it constitutes a fishery of considerable importance in the small-scale sector while the fidelity of the species to certain regions/areas together with the growing industrial and ever-expanding artisanal fleets in such areas is looming largely on the sustainability of the fisheries. IOTC is the RFMO monitoring and managing the fishery for the species in the Indian Ocean while in the West Central Pacific (WCP), the species is managed by the individual countries (Willette *et al.,*

2016). The IOTC (2017a) reported the status of the stock to be overfished and subject to overfishing while that in the WCP region *per se* is not known though many nations reported decline in catches However, the IOTC's management advice for the region was with a view of uncertainness of stock structure and the total catches and it called for collaborative work to delineate the stock for better assessment of the stock in the larger Indian Ocean basin.

Connectivity of the LOT throughout its distribution is one of the concerns for global management of the resource. Recent study (Willette *et al*., 2016) hinted discrete genetic stocks between Indian Ocean and the South China Sea, the two major shelves of its dominance while ruling out any stock structure in the individual basins (Kunal *et al.,* 2014 and Willette *et al*., 2016). Molecular studies covering the northern Indian Ocean along the coasts of Oman, Iran and Pakistan and the coasts along the Bay of Bengal and Andaman Sea together with the western coasts of Indonesia and Australia would expand the horizon of the knowledge on population structure in the larger Indian Ocean basin further adding to the connectivity with the adjacent basins. In an effort in this direction, the present study attempts to connect the South China Sea and the Northern Indian Ocean shelves by studying the samples from Kochi in the southern west coast of India along the eastern AS and Port Blair along the western Andaman Sea together with additional samples from the northwest coast.

Materials and Methods

Sample collection, DNA extraction and amplification of the two mitochondrial fragments

Tissues sample of LOT was collected during 2014 and 2015 from 3 different locations: Veraval, Kochi and Andaman Islands (Fig.1). DNA extraction was carried out using standard protocols. Partial control region and COI gene of *Thunnus tonggol* was amplified using universal primers (like in Cheng *et al.*, 2012; Folmer *et al*., 1994). PCR reactions were carried out in 25 µl reaction mixture containing 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 200 µM of each dNTP, 0.2 µM of each oligonucleotide, 1 unit of Taq DNA polymerase and 50 ng of template DNA using Biorad T100 thermocycler (Biorad, USA). PCR programme consisted of an initial denaturation at 94^0C for 4 minutes, 35 and 33 cycles of denaturation (for control region and COI respectively) at 94^0C for 30S, annealing at 48^0C and 42^0C (for control region and COI respectively) for 30S, extension at 72^0C for 55S and 40S (for control region and COI respectively) and a final extension at 72^0C for 7 min. The PCR products were purified using Qiagen PCR purification kit and sequencing was carried out using the BigDye Terminator

Sequencing Ready Reaction v3.0 kit (Applied Biosystems) on an ABI 3730 automated sequencer. A 758bp region of the control region and a 576bp region of COI of LOT was amplified in all the sampled individuals for phylogenetic and population genetic analysis. All the partial sequences of LOT control region and COI were deposited in GenBank with the accession nos. (MF592988-MF593048; MG720822-MG720850). In addition to the sequences generated in the present study, sequences deposited in GenBank from Ratnagiri and Veraval coasts of India for control region were also included in the analyses.

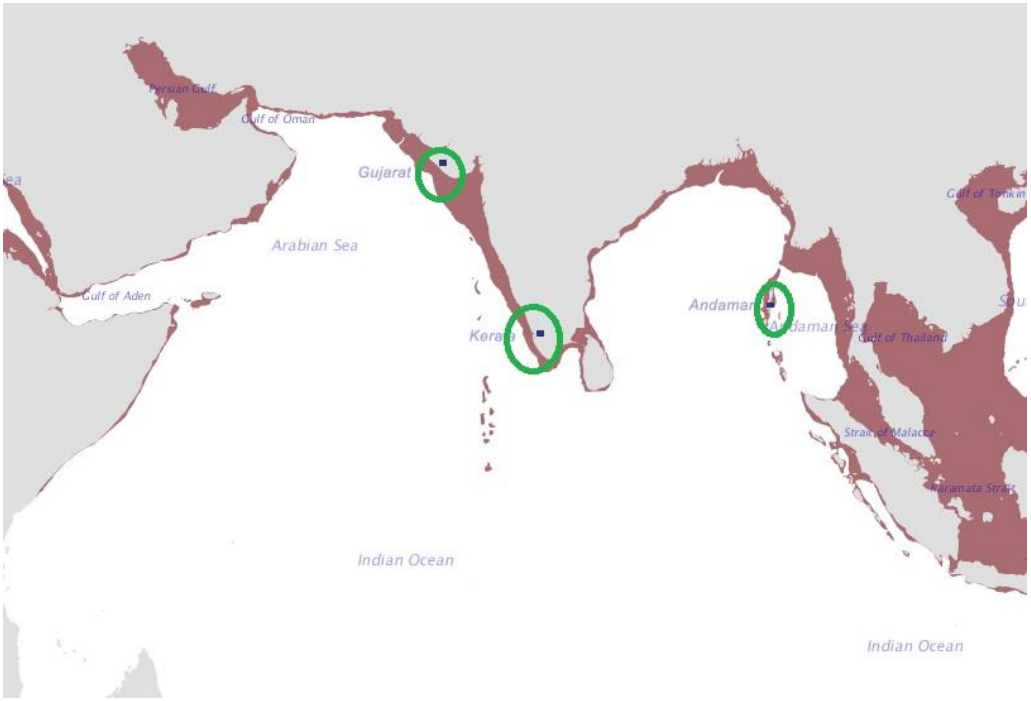

Figure 5.1: Sampling locations for stock structure studies of LOT.

Results

The analysis of 853bp portion of the LOT control region excluding gaps showed very high levels of variation. A total of 61 individuals were collected from Veraval, Kochi and Andamans in addition to the 92 haplotypes retrieved from NCBI (those of Veraval and Ratnagiri) of a previous study from the Indian Ocean region (total number of analysed sequences, 153). 101 variable sites consisted of 22 singleton sites and 79 parsimony informative sites with average value of nucleotide differences being 13.42. Overall haplotype (Hd) and nucleotide diversity (π) for all sequences were 0.998 and 0.037 respectively. COI gene region was analysed for a portion of 622bp and a total of 9 haplotypes were observed from 29 individuals. There were 8 variable sites consisting of 5 singleton sites and 3 parsimony informative sites with average value of nucleotide differences being 0.921. Overall haplotype

(Hd) and nucleotide diversity (π) for all sequences were 0.53 and 0.0014 respectively. The basic statistics of both sequence sets are presented in Table 1 and 2.

Table 1: Summary of analysis of molecular variance (AMOVA) carried out using mitochondrial control region sequences

Structure tested	Observed partition			
	Variance	%total	Φ statistics	P (after Bonferroni correction)
1. One gene pool (Gujarat, Kochi,Andaman)				
Among populations	0.17	1.69	Φ_{ST} = 0.016	>0.05
Within populations	9.99	98.31		
2. Two gene pools(Gujarat, Kochi) (Andaman)				
Among groups	0.39	3.84	Φ_{CT} = 0.038	>0.05
Within groups	-0.045	-0.43	Φ_{SC} = -0.004	>0.05
Within populations			Φ_{ST} = 0.034	>0.05

Table 2: Summary of analysis of molecular variance (AMOVA) carried out using mitochondrial cytochrome c oxidase1 sequences

Structure tested	Observed partition			
	Variance	%total	Φ statistics	P (after Bonferroni correction)
1. One gene pool (Gujarat, Kochi,Andaman)				
Among populations	0.032	6.82	Φ_{ST} =0.068	>0.05
Within populations	0.438	93.18		
2. Two gene pools(Gujarat, Kochi) (Andaman)				
Among groups	-0.019	-4.22	Φ_{CT} =	>0.05

Within groups	0.044	9.78	Φ_{SC} =	>0.05
Within populations	0.44	94.44	Φ_{ST} =	>0.05

The best evolutionary model for nucleotide substitution for the control region data was the Tamura 3-parameter (T92) (Tamura, 1992) with invariable sites (I=0.60) and rate heterogeneity (G=0.61) (T92+G+I) and for COI gene region nucleotide substitution was Kimura 2p. The phylogenetic tree constructed using both sets of sequences supported the existence of two well separated clades or haplogroups with significant bootstrap values without any geographic pattern (Fig. 2). Similar pattern of two well separated clades was observed in haplotype network diagram constructed using both sets of sequences (Fig.3). Genetic differentiation among LOT populations from different locations was tested using ΦST pair-wise comparisons of both control region sequences and COI gene sequences. There were 9 possible comparisons and none of them were statistically significant after Bonferroni correction. The results were further confirmed by a hierarchical AMOVA test by partitioning the variance among and within populations i) one gene pool comparison considering all the populations as a single gene pool, ii) among groups, within groups and within populations: two gene pools [Mainland India; (Veraval & Kochi) versus Andaman] and three gene pools (Veraval versus Kochi versus Andaman). The overall levels of genetic differentiation were not significant ($p > 0.001$) in all the analyses with a global ΦST value of 0.0.017 with control region sequences and 0.068 with COI gene sequences.

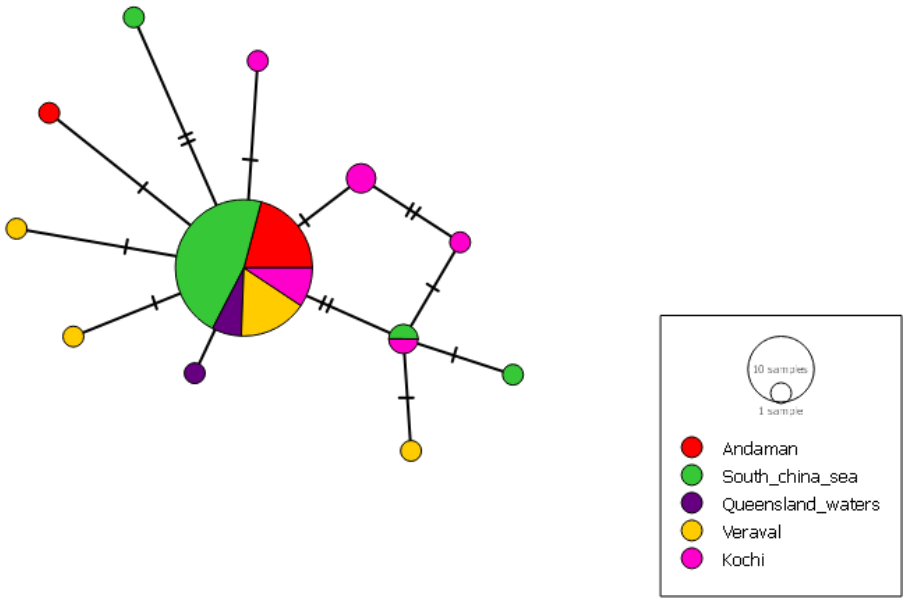

Figure 2: Median joining network diagram constructed using COI sequences

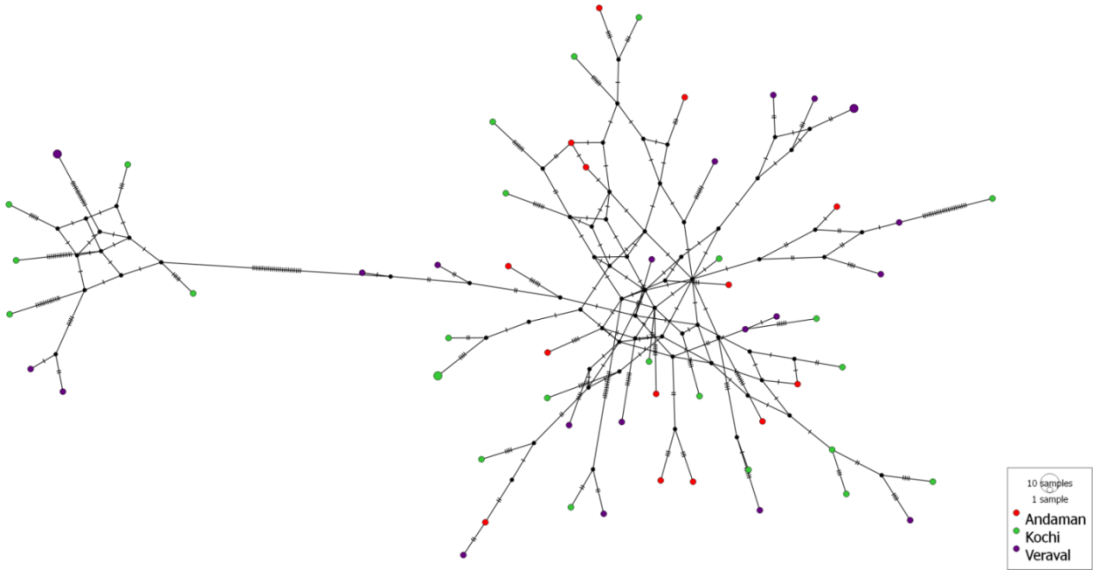

Figure 3: Median joining network diagram constructed using Control region sequences

The mismatch distribution of both control region sequences, and COI gene sequences showed bimodal pattern when all samples were included (Fig. 4) indicating a stable population history. Tajima's D values and Fu's Fs (-1.67 and -1.01 respectively) were not

significant (p>0.05) indicating absence of population expansion or bottleneck. There was no evidence of isolation by distance as there was a lack of significant correlation (*p*>0.05) between FSTand (1-FST) and geographic distance.

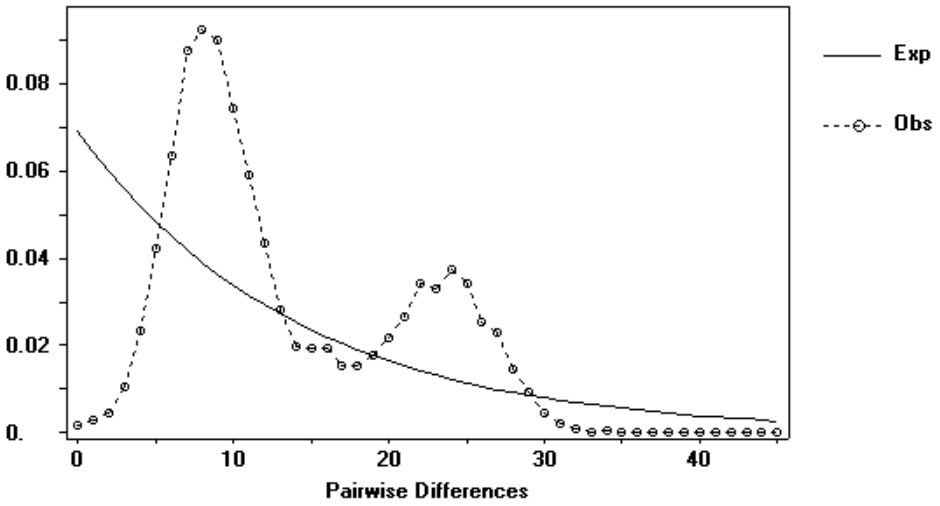

Figure 4**:** Mismatch analysis plot constructed using the whole set of sequences

Discussion

Stock Structure

Single locus mtDNA studies continued to provide powerful first assessment of phylogeographic patterns of otherwise complex population structure of marine organisms (Bowel *et al.,* 2014). LOT, a neritic tuna species has unique distribution and abundance pattern spanning the shelves of Indo-Pacific with certain regions in the Indian Ocean dominating the fishery. Present study was intended to check if there are multiple stocks of LOT in the wider Indian Ocean through sampling from geographically well separated (three) regions, one in Gujarat, the northern end of west coast of India and Kerala, the Southern end set apart by nearly 10^0 latitudes (over 1000 km), and the Andaman & Nicobar Islands in the Andaman Sea set apart from the other two regions by nearly 20^o longitudes (nearly 4000 km). Utilization of the published sequence data (Kunal *et al.,* 2014) from Ratnagiri in the central west coast of India enhanced the scope of the study. The results were confirmed further by analyzing the COI sequences of samples from these regions as well as that from the South China Sea and Queensland waters (retrieved from GenBank) thus extending the scope further.

The results ruled out existence of any stock structure of LOT in the wider northern Indian Ocean region along the Indian peninsula as well as the archipelagic waters of Andaman & Nicobar Islands though two haplogroups *sans* any geographic pattern were observed in the eastern AS along the west coast of India. The outcome is supported by results of Phylogenetic tree and AMOVA. Similar observations were also made by Kunal e*t al.* (2014) based on their study covering two sampling points along the northwest coast of India asserting the existence of single panmictic population of LOT in the northwestern shelf of India. Analysis of COI sequences while confirming the homogeneity of the stock in the Indian Ocean signaled connectivity of these stock with that of the West Central Pacific.

Tunas in general are known to have homogenous population within and between the oceans (Alvarado-Bremer *et al.,* 1998; Grewe & Hampton, 1998; Chow *et al.,* 2000; Appleyard *et al.,* 2002; Durand *et al.,* 2005; Ely *et al.,* 2005, Chiang *et al.,* 2008) because of the occurrence of continuous, circumtropical pelagic environment and a wide range of suitable spawning grounds (Kumar *et al*., 2012a). Lack of genetic structure demonstrated extensive gene flow within and across the oceans (Chiang *et al.*, 2008). Coastal pelagic fishes generally have more geographic population subdivisions than do oceanic species (Crosetti *et al.,* 1994; Rossi *et al*. (1998) probably because of their inshore habitat requirements, shorter migration distances, and vulnerability to climatic fluctuations. *Auxis thazard,* a smaller neritic tuna has three distinct stocks in the Indian Ocean one around the Andaman Sea, second around southwest coast of India, and the third around rest of the Indian coastal waters (Kumar *et al.*, 2012a). But, several neritic species of tuna and tuna like fishes are reported to have genetically homogenous populations in different oceans like the kawakawa in the Indian Ocean (Kumar *et al.*, 2012b) and South China Sea (Santos *et al.*, 2010), LOT along the coast of India (Kunal *et al.,* 2014), narrow barred spanish mackerel (*Scomberomorus commerson*) in the ROPME Sea areas (Hoolihan, 2006), Atlantic spanish mackerel (*S maculate*) in the western Atlantic Ocean and Gulf of Mexico (Buonaccorsi *et al.,* 1999), Japanese spanish mackerel (*S niphonius*) in the East China Sea and the Yellow Sea (Shui *et al.*, 2009). *Auxis spp* is characterized by faster growth, shorter life span and maximum reported growth of 57.95cm (Ghosh *et al.*, 2012) and is known to undertake shorter migrations which may be the reason for existence of multiple stock of this species (Kumar *et al.*, 2012c). Whereas *T tonggol* being a slow growing and long lived species (Yesaki, 1993; Griffith *et al.,* 2010a) was reported to have grown up to 145.2cm (Dawood *et al.*, 2014).

Continuity of the shelf areas and absence of physical barrier in the entire northern Indian Ocean region in general and the shelf of mainland India in particular facilitate free movement of adults and larval stages of the fishes. Monsoon influenced currents in the Indian Ocean (Schott and Mc Creary, 2001) aids dispersal of the eggs and larvae to wider areas causing mixing of the populations along the coast and across the basins. Tropical Indian Ocean exhibits circulation patterns drastically different from Atlantic and Pacific oceans as it is affected by monsoon winds. Two westward flowing currents (North Equatorial currents and South Equatorial currents) and one eastward flowing current (Equatorial Counter Current-ECC) persist during the noutheast monsoon season. The flow gets reversed eastward during the southwest monsoon season and combines with ECC which is called southwest Monsoon Current (SMC) (Varkey *et al.,* 1996). Substantial mixing of Bay of Bengal and AS waters occur during these currents which causes dispersal of pelagic larvae. The Indonesian Through flow (ITF), a wind-driven circulation connects the Pacific Ocean with the Indian and Atlantic oceans via the Indonesian passages and the Agulhas and Tasman gateways (Ridgway & Dunn, 2007; Speich *et al*., 2007; Lambert *et al.,* 2016).

Oceanic tunas mostly have multiple stocks in all the major oceans. Skipjack (*Katsuwonus pelamis*), a circum global tropical oceanic tuna is said to have four discrete stocks in the Indian Ocean (Menezes *et al.,* 2012) and the Atlantic bluefin tuna (*T thynnus*) have two separate spawning stock in the northern Atlantic (Rooker *et al*., 2007). Bluefin tuna have discrete stocks in Pacific Ocean (Ward *et al*., 1995) and Mediterranean Sea (Carlson *et al*., 2004 and Riccioni *et al.,* 2010), and Carlsson *et al.,* (2007) and Boustany *et al.* (2008) reported the stocks between Atlantic Ocean and Mediterranean Sea to be heterogeneous. Later studies have revealed that the Yellowfin tuna has discrete populations in the Indian Ocean (Dammangodda, 2008) and along the coastal waters of India (Kunal *et al.,* 2013). Ward *et al.*, 1997 reported four discrete stocks one each in the Indian Ocean, Atlantic Ocean, western Pacific Ocean and eastern Pacific Ocean. The Atlantic blackfin tuna (*Thunnus atlanticus*), the smallest species of the *Thunnus* genera has discrete stocks in Gulf of Mexico and the western Atlantic Ocean (Saxton, 2009). However, though the information is scanty, the bigeye tuna is an exception to the other oceanic tuna as it has single panmictic population within the Indian Ocean (Appleyard *et al.*, 2002) and between the Indian Ocean and the West Pacific Ocean (Chiang *et al*., 2008).

Previous study on the genetic diversity of LOT in the Indian Ocean region by Kunal *et al.* (2014) also reported existence of two clades along the coast of India without any geographic pattern however; the species didn't show any genetic variation of that kind in South China Sea (Willette *et al.,* 2016). Such genetic discontinuities were reported for kawakawa, a neritic species in the South China Sea (Santos *et al.,* 2010) and bigeye tuna, an oceanic species in the Atlantic, Indian and Pacific Oceans (Chiang *et al.,* 2008). As opined in most cases (Kunal *et al.,* 2014; Santos *et al.,* 2010; Chiang *et al.*, 2008; Kumar *et al.,* 2012b), the observed genetic mitochondrial discontinuities may have originated by common vicariant events resulting from a general lowering of the temperature of the water that produced reduction of tropical marine habitats, and isolation of populations during Pleistocene glacial maxima (Alvarado Bremer *et al.,* 1998; Graves & McDowell, 2003; Viñas *et al.,* 2004; Alvarado Bremer *et al.,* 2005) followed by a secondary contact during inter-glacial periods by unidirectional gene flow of formerly allopatric populations (Peeters *et al.,* 2004; Alvarado Bremer *et al.,* 2005). Observation by Nei, (1987) and Avise, 1998 that the characteristic large population sizes, environmental heterogeneity and life-history traits which favour rapid population increase as the explanation for high haplotypic diversity within populations doesn't hold good in this case as there are no drastic variations in the environment in the distributional area of the species along the coasts of India and the species is relatively slow growing and long lived (Griffith *et al.,* 2010a). Non-existence of genetic diversity in the Andaman Sea and the West Pacific as in the western coast of India indicates poor mixing of the population from these the geographically distinct regions or that the gene flow is unidirectional from the Indo-Pacific to the Indian Ocean as suggested by Durand *et al.* (2005) and Chiang *et al.* (2008) in the case of bigeye tuna. This also suggests existence of a self-recruiting population of the species in the Andaman coasts.

Knowledge on the biology and ecology of the study organisms greatly enhances the phylogeogrpahic inferences (Bowen *et al*., 2014). Spatial and temporal information on spawning and early life history all along the distributional range of the species would elucidate if there are discrete spawning populations (Neilson *et al.,* 1988). However, information on the ecology, spawning and early life history is scanty for the LOT throughout its worldwide distribution. The *T tonggol* is unique among the tunas in its distributional pattern of exclusivity to neritic areas close to the land masses. They are rarely found offshore and markedly absent in the oceanic islands like Maldives and Lakshadweep (Yesaki, 1993) unlike other neritic

species like *E affinis* and *Auxis* spp. The species tend to avoid areas near the mouths of larger estuaries with low salinity and high turbidity (Collette & Nauen, 1983). It's a warm water species preferring the temperature regime of 28-30^{o}C (Yesaki, 1993). As per the available information, the spawning grounds for this species exist in the Gulf of Thailand (Yesaki, 1982) and western Sea of Japan and the East China Sea (Itoh *et al.*, 1999). There isn't any clear information on the spawning and early life history of this species in the northern Indian Ocean where the species enjoys considerable abundance. Length frequency studies from the Indian peninsula covering southwest coast of India to the coasts of Oman as discussed by Abdussamad *et al*. (2012a) and the biology studied in this thesis (as discussed in chapter 3) indicated an ontogenetic migration of the species towards northern latitude as they grow. Similar movements were observed along the coasts of Australia, based on which Griffith *et al*. (2010) suggested the nursery ground to be along the northern Australia (which was partly corroborated by the studies on the reproductive biology mentioned therein). Reports on the reproductive biology of the species along the northwest coast of India as discussed in this thesis under chapter 3 substantiated the spawning ground to be the outer neritic areas of the northern Arabian Sea/Oman Sea close to Oman and the spawning peaked around the month of May. However, juvenile fishes of the size less than 25cm were reported from southwest coast (Kerala and Karnataka) (Muthiah, 1986 & Abdussamad *et al*., 2012a), northwest coast (Gujarat) (Abdussamad *et al*., 2012a and chapter 3 of this thesis*)* indicates possibility of nursery grounds in these regions and hence occurrence of multiple spawning locations in the outer neritic areas along the west coast of India. There is an absolute lack of information on the biology of the fish from the Andaman Sea or Bay of Bengal. However, considering the distance from fishing grounds off the Indian mainland, the Andaman Sea is likely to have self-recruiting population of *T.tonggol.*

When the stocks are truly panmictic, managing the fishery as a single stock will not affect the recruitment from or to overfished areas but if however, there are different populations, management as a single stock will mean that over-exploitation in particular areas will lead to reductions in effective population size and yield in those areas (Chiang et al., 2008). In addition, if components of stock complex exhibit high levels of gene flow, then management should not be based on genetic data alone (Waples, 1998). Other data such as tagging, morphology, otolith chemistry, life histories and catch rates are required to help determine management units. Future study on the stock structure of this species in this region

need focus on wider coverage of sampling right from its western extreme in the Indian Ocean along the coasts of Africa to both coasts of India, Sri Lanka, Andaman, west coast of Thailand to up to the southwestern Australia, preferably with the advanced markers in combination with the otolith chemistry and other methods as a collaborative activity of the interested nations.