Analysis of genetic variation in green chromide \([\text{Etroplus suratensis} \text{ (Bloch)}]\) (Pisces: Cichlidae) using microsatellites and mitochondrial DNA

M Dhanya Alex\(^1\), A Biju Kumar\(^*\), U Suresh Kumar\(^2\) and Sanil George\(^2\)

\(^1\)Department of Aquatic Biology and Fisheries, University of Kerala, Thiruvananthapuram 695 581, India
\(^2\)Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram 695 014, India

Received 5 May 2015; revised 27 May 2015; accepted 4 June 2015

The cichlid fish, Green Chromide \([\text{Etroplus suratensis} \text{ (Bloch)}]\) is an economically valuable food fish and a preferred candidate for brackishwater aquaculture in India. Genetic variation of \(E. \text{suratensis}\) collected from 11 different geographic locations of Kerala state, India was investigated using microsatellite and mitochondrial DNA markers \([\text{cytochrome c oxidase 1 (CO1)} \text{ and } 16S \text{ rDNA genes}]\). Seventeen primers published for two cichlid species were tested to amplify homologous microsatellite loci in \(E. \text{suratensis}\). Six primers yielded successful amplification and only one was found to be polymorphic. Microsatellite analysis revealed a low genetic variation with alleles ranging from 1 to 4. The observed heterozygosity and expected heterozygosity ranged from 0.000 to 0.167 and 0.000 to 0.104. The results showed relatively lower variation in Kayamkulam and Vembanad lake populations. Mitochondrial DNA (CO1) analysis revealed 9 haplotypes with very low haplotype and nucleotide diversity. Genetic differentiation by pairwise \(F_{ST}\) showed that samples from Chandragiripuzha estuary are significantly different from all other populations. AMOVA analysis of both markers indicates significant differentiation between populations. The results of the two markers suggest that the studied 11 populations of \(E. \text{suratensis}\) showed low genetic variation between populations and leads to the conclusion that they are drawn from the same randomly mating gene pool.

\textbf{Keywords:} genetic variation, heterozygosity, microsatellite DNA, mitochondrial DNA, Pearl Spot, population

Introduction

Cichlidae, a species-rich family of perciform fishes, currently enjoying Gondwanan distribution, has attracted much attention of fishery scientists and evolutionary biologists. While other taxa of cichlids are found in Africa, the monophyletic genus \(\text{Etroplus} \text{ G. Cuvier, 1830}\) is found naturally in inland waters of southern India and Sri Lanka\(^1\). The Green Chromide or Pearl Spot, \(\text{Etroplus suratensis} \text{ (Bloch)}\) is naturally distributed in South India and Sri Lanka\(^2\), and is a preferred candidate for brackishwater aquaculture in India because of high demand as food fish.

Genetic data have been increasingly used for the conservation and management of endangered species. Microsatellites or simple sequence repeats (SSRs) are proven tools for direct assessment of genetic variation and population level evolution\(^3\). Mitochondrial DNA (mtDNA) is now emerging as the most preferred choice for assessing both population structure and genetic variability, and useful for investigating phylogeographic groups within a single species\(^4\). Despite the importance of \(E. \text{suratensis}\) as a highly valuable fishery resource in India and a potential candidate for brackishwater aquaculture, there is no information on the genetic variation of this species, which has been declared as the state fish of Kerala. The present paper documents the genetic variation of \(E. \text{suratensis}\) populations inhabiting various aquatic ecosystems of Kerala state, India using microsatellite and mitochondrial DNA (CO1 & 16S rDNA) markers.

Materials and Methods

Sampling and DNA Extraction

In total, 11 geographically isolated locations of \(E. \text{suratensis}\) comprising various rivers and lakes of Kerala were selected for the sampling (Fig. 1). The collection sites were Poovar estuary, Vellayani lake, Veli lake, Ashtamudi lake, Kayamkulam lake, Vembanad lake, aquaculture pond at Regional Agricultural Research Station (RARS) Kumarakom, Chalakkudy river, Biyyam kayal, Malampuzha reservoir and Chandragiripuzha estuary. Genomic DNA was isolated from ethanol preserved tissue
samples by phenol-chloroform method with minor modifications. The quality and concentration of DNA samples were assessed by 0.8% agarose gel electrophoresis and then the samples were stored at −20°C for further use.

Molecular Analysis

Microsatellite Analysis

For microsatellite analysis, 110 specimens of *E. suratensis* from 11 populations were selected. A total of 17 primers were used in the cross-species amplification of microsatellite loci for population genetic analysis. PCR amplification was performed on a thermal cycler (Applied Biosystems) and the PCR program was as follows: 30 sec at 98°C, followed by 30 cycles of 5 sec at 98°C, 10 sec at the annealing temperature and 15 sec at 72°C, with a final extension of 1 min at 72°C. PCR products were separated on an ABI 3500 Genetic Analyzer and analyzed using GeneMapper ID-X v1.4 software.

PowerMarker version 3.23 was used to calculate the number of alleles per locus, gene diversity and the polymorphic information content (PIC). Genetic diversity parameters, such as, major allele frequency, observed number of alleles (na), effective number of alleles (ne), observed heterozygosity (Ho), expected heterozygosity (He) and Nei’s (1978) unbiased genetic identity and distance were estimated using the program GenAlEx.

Mitochondrial DNA Analysis

Amplifications of CO1 gene were carried out using two primers: Fish F1 (5'-TCA ACC AAC CAC AAA GAC ATT GGC AC-3') and Fish R1 (5'-TAG ACT TCT GGG TGG CCA AAG AAT CA-3'). PCR amplification reactions were run in a PCR thermal cycler (Gene Amp PCR system 9700, Applied Biosystems) in the following conditions: preliminary denaturation at 95°C for 5 min, followed by 10 cycles consisting of denaturation at 95°C for 30 sec, primer annealing at 45°C for 40 sec, primer extension at 72°C for 1 min and also followed by another 30 cycles of denaturation at 95°C for 30 sec, primer annealing at 50°C for 40 sec, primer extension at 72°C for 1 min, and concluded by a final extension step at 72°C for 10 min.

The partial sequence of 16S ribosomal DNA was amplified for 33 individuals of *E. suratensis* using primers 16S FP (5’CGCCTGTATCAAAAACAT3’) and 16S RP (5’GTCTGACTCAGCTACGT3’). Amplification conditions were as follows: 98°C for 30 sec; 40 cycles at 98°C for 5 sec, 48°C for 10 sec and 72°C for 15 sec, and a final extension at 72°C for 1 min.

The sequence quality was checked using sequence scanner software V1 (Applied Biosystems) and DNA sequences were aligned using Geneious Pro v 5.1 software package. The number of haplotypes, haplotype diversity and nucleotide diversity were calculated using the program DnaSP version 3.0. Phylogenetic and molecular evolutionary analysis was conducted using MEGA version 4. Sequences of the mtDNA CO1 and 16S region of *E. suratensis* were deposited in GenBank, under the accession numbers KF442162 to KF442194 and KF442129 to KF442161.

Results

After cross-species amplification of 17 primers from 2 cichlid fishes, 6 primers yielded amplified products. Only one microsatellite locus, OM08 from *Oreochromis mossambicus*, was polymorphic and this was selected for genetic variation analysis. A total of 9 alleles were detected, ranging in size from 96 bp to 384 bp. Only one allele was observed in all the five loci, except OM08. Allele frequencies of populations over six loci are represented in Fig. 1. The mean number of alleles per locus per population was 1.242. The Poovar population had the highest mean number of alleles with 1.5 alleles, followed by Chalakkudy, Biyyam Kayal, Malampuzha and Chandragiripuzha with 1.3 alleles; Vellayani, Veli, Ashtamudi, Kumarakom and Vembanad with 1.2 alleles; and Kayamkulam population with only one allele. The number of the effective alleles per locus ranged from 1.000 to 1.278 with an average of 1.103. Highest effective number of alleles was observed in Poovar population. Observed heterozygosity (Ho) ranged...
from 0.000 to 0.167 and expected heterozygosity (He) ranged from 0.000 to 0.104. The average Ho (0.070) and the He (0.650) were calculated at the population level.

Genetic differentiation between populations was analysed using $F_{ST}$ pairwise differences. The largest genetic differentiation between populations was between Vellayani and Kayamkulam populations. The coefficient of genetic differentiation values ranged from 0.004 (between Ashtamudi lake & Vembanad lake, and between Biyyam Kayal & Chalakkudy river) to 0.241 (between Vellayani lake & Kayamkulam lake). The AMOVA analysis with consideration between all 11 populations revealed that almost all of the variance in data, namely, 81% (p=0.001), was within populations and genetic variances among populations was only 19% (p=0.001). Nei’s genetic distance and identity estimated between pairs of 11 populations of *E. suratensis* showed that highest genetic distance between Poovar estuary and Kayamkulam populations, and lowest genetic distance between Ashtamudi lake and Vembanad lake, and between Kayamkulam lake and Vembanad lake populations. Similarly, higher genetic identity was recorded between Ashtamudi lake and Kayamkulam lake populations. Molecular diversity indices showed that the average gene diversity value was 0.0643.

Cluster analysis of 11 populations using the UPGMA approach revealed 4 major groups (Fig. 2), indicating the locations of *E. suratensis* populations which were nearer and grouped in one cluster by dendrograms, except two populations (Poovar estuary & Chandragiripuzha estuary).

Mitochondrial CO1 gene of 700 bp size was amplified and sequenced from 33 specimens of *E. suratensis* and was used to study the genetic variation among 11 different geographic locations of Kerala. CO1 gene revealed 11 variable and 4 parsomonic informative sites in 700 bp long region. The mean total nucleotide composition was A=23.39%, T=22.1%, C=17.85% and G=28.42%. A total of 9 haplotypes were identified among 11 populations of *E. suratensis*. Among the 9 haplotypes recognized, the most widespread haplotype (haplotype 1) was found in all populations, except in Chandragiripuzha estuarine population. Two haplotypes were observed in Chandragiripuzha estuarine population (haplotype 8 & 9). Specific haplotypes were observed within populations of Veli, Kayamkulam, Kumarakom and Vembanad. Haplotype diversity among all 11 populations was 0.47538, whereas nucleotide diversity was 0.00136. Highest nucleotide diversity was observed in Chandragiripuzha population (0.00509).

Genetic distance between the populations ranged from 0.0000 to 0.0052, with the highest value between Chalakkudy river and Chandragiripuzha estuarine populations (0.0052). Genetic differentiation by pairwise $F_{ST}$ showed that samples from Chandragiripuzha estuary are significantly diverged from all other populations. No genetic differentiation was observed in any other pairs of populations. Population pairwise $F_{ST}$ value ranged from 0.0000 to 0.4444. The hierarchical AMOVA of population structure reveals a highly significant subdivision between populations. The $F_{ST}$ value for all the 11 populations was found to be 0.267 (P<0.05). AMOVA results indicated that 26.67% of the total variation existed among populations and 73.33% of the variation within populations. Phylogenetic trees constructed using the Maximum Parsimony (MP) and UPGMA methods (Fig. 3) show the same topology and consistently displayed that all samples from the 11 populations were mixed with one another and were clustered into a single group, except two samples from Chandragiripuzha estuarine population, indicating that almost all samples of *E. suratensis* from different geographic locations of Kerala did not
form obvious bunch groups according to the spatial distribution and revealed the migration from one population to another.

Comparison of the 575 bp sequence of 16S rDNA region revealed the presence of only two haplotypes out of 33 samples of _E. suratensis_. Sequence variation in the 16S was very less extensive than CO1 gene regions. No parsimony informative sites were found and the mean nucleotide composition was A=29.7%, T=30.34%, C=26.3% and G=21.9%. Only two haplotypes were detected and it was clear that the haplotype H1 was widespread in all populations and haplotype H2 was observed in one sample of Veli lake population. Nucleotide diversity and haplotype diversity of Veli lake population was 0.00116 and 0.66667, respectively; while the other populations were zero because of no variable sites in it. Nucleotide diversity in entire population was observed as 0.00011, whereas the haplotype diversity was 0.06061. As observed in CO1, phylogenetic analysis of 16S rDNA also indicated that almost all samples of _E. suratensis_ did not form obvious cluster according to the spatial distribution, except one sample from the Veli population which is clustered into a single group.

**Discussion**

Despite the importance of _E. suratensis_ as a highly valuable fishery resource in Kerala, there is little information regarding the genetic background of this species. The present study is the first report on the genetic status of this species using microsatellite and mitochondrial (CO1 & 16S rDNA) markers.

Microsatellites are highly conservative in their flanking regions and hence can persist for long evolutionary time spans much unchanged. Due to this, primers developed for a species from the flanking regions of a microsatellite locus can be used to amplify the same locus in other related species. In the present study, successful cross priming was obtained with 6 primer pairs; of which 5 primers were monomorphic and only one was of polymorphic in nature, which was ideal to be used as marker in stock identification studies of _E. suratensis_. However, the optimum annealing temperature to get scorable band in _E. suratensis_ slightly differed from that reported for the respective primer pair in the resource species.

The number of alleles at different microsatellite loci in _E. suratensis_ varied from 1 to 4 with an average value of 3. Similar values of allelic variation was recorded in _Tor tambroides_ with number of alleles ranged from 2 to 5 and _T. douronensis_ (2 to 5 alleles/locus). The reason in low level of allele variation is possibly due to small sample size. Previous reports showed that marine species have greater microsatellite allele variation as compared with freshwater species and this was consistent with the increased evolutionary population sizes of marine species. They also reported that much of the variation in polymorphism at microsatellite loci that exist between species and classes can be attributed to differences in population biology and life history, and to a lesser extent to differences in natural selection pertaining to the function of the microsatellite loci.

The mean Ho and He were relatively lower in _E. suratensis_ compared to those in the previous reports. Inbreeding and non-random mating would
result in heterozygote deficit. Over exploitation may lead to reduction in effective population size of fish species across the world and *E. suratensis* is currently an over exploited species in Kerala. In overexploited populations, inbreeding can happen that may result in deficiency of heterozygotes and this could be considered as the reason for low heterozygosity observed in *E. suratensis* inhabiting various water bodies of Kerala.

Relatively lower variation in Kayamkulam lake and Vembanad lake populations suggests that these populations may be of more concern from a conservation viewpoint. Previous studies reported the consistent reduction of *E. suratensis* landings from the wetlands of Kerala, especially, from Vembanad lake due to ecological degradation. Relatively low genetic variation could have been a result of population isolation, small population size or historical population bottleneck. Small genetic variation can result in decline of adaptability to changing environments.

In summary, microsatellite analysis showed relatively low genetic variation, which may impair adaptability to the changing environment. Therefore measures should be established to avoid or minimize the additional threats caused by anthropogenic activities on these populations, such as, establishment of protected areas, restriction of fishing sizes, fishing ban during spawning aggregation period etc. The current microsatellite analysis forms the baseline information on the genetic variation of the state fish of Kerala. However, further studies using more number of polymorphic primers should be conducted to corroborate the present findings.

In mitochondrial analysis, out of the two genes analysed, COI exhibited significantly higher genetic divergence than 16S rDNA as expected with higher rate of evolution. Nine haplotypes were observed in the present study among 33 individuals from 11 populations of *E. suratensis*. Similar results were reported in *E. suratensis* with 7 haplotypes. The study reports that haplotype 1 was found in all populations, except for Chandragiripuzha estuarine population as it might be the evidence of its isolation. The two haplotypes observed through 16S rRNA in the current study can attribute to the findings of Ayres et al. who suggested that all populations were founded on individuals from a common source.

Intra-population nucleotide diversity and haplotype diversity was very low in *E. suratensis*. These findings are supported by the previous findings reported in Malaysian Borneo, *T. tambroides*. The low intra-population haplotype diversity in the present study may be due to the low population size. Low nucleotide diversity values are also indicative of low population diversity and genetic diversity influenced by many factors, such as, anthropogenic activity and habitation, bottleneck effects and founder effects. Very low or no nucleotide diversity observed in Veli lake, Kayamkulam lake and Vembanad lake might be due to the smaller population of *E. suratensis* available in these lakes, which are in ecological debilitation due to overfishing, reclamation and pollution.

The results of AMOVA were found to be significant for the genetic structuring of 11 populations and the overall *F*<sub>ST</sub> value (0.27; P<0.05) for all the populations was found significant. AMOVA results are in contrary with the previous findings in *E. suratensis* populations collected from Kerala, Tamil Nadu and Maharashtra, wherein 42.05% of variation among populations and 73.33% within populations were reported. Population differentiation depends directly on gene flow. Species with higher capabilities of dispersal and migration across geographic barriers do not show much population differentiation, whereas other species with lower dispersal and migration capabilities exhibit significant differentiation among populations over time.

Genetic differentiation by pair-wise *F*<sub>ST</sub> analysis indicated that samples from Chandragiripuzha estuarine population significantly diverged from all other populations. No genetic differentiation was observed in any other pairs of populations. Isolation by distance and genetic drift might be the main effect that may explain the genetic divergence of Chandragiripuzha estuarine population, since it is geographically distant from all other sampling sites and considered one of the sites that is nearer to Karnataka state and North to the Palghat Gap. The distinction of the Chandragiripuzha estuarine population from the rest of the populations of *E. suratensis* is supported by microsatellite analysis of the present study. Another explanation for the lack of isolation between the populations might be the translocation of *E. suratensis* for aquaculture and trade, and ranching programme in rivers and lakes initiated by the government. According to previous reports geographic segregation, habitat differences
and a sudden reduction in a colony resulted in increased divergence, whereas migration and human activities could decrease divergence.\(^{18,24}\)

As reported in the microsatellite analysis of the present study, Chandragiripuruzha estuarine population deviated as separate clade in the dendrogram of CO1 because this population is geographically distinct from other populations and geographically placed towards northern most part of Kerala, close to the Karnataka state. Since CO1 gene has low variation and highly conserved, the differentiation seen in one sample of Veli lake is insignificant. The distinction of one sample of Veli lake population in 16S rDNA is unsubstantial as this gene is highly conserved and accordingly very low variation compared to CO1.

In light of the above, the present findings using two markers, microsatellite and mitochondrial DNA (CO1 & 16S rRNA), suggest that the studied 11 populations of *E. suratensis* inhabiting different geographic locations of Kerala show low genetic variation between populations because they are drawn from the same randomly mating gene pool. Both the markers revealed a remarkably high genetic differentiation among the populations indicating population fragmentation and reduction of gene pool. The loss of genetic variation among populations might be due to the fact that populations are highly differentiated with one another. This can occur when physical barriers are removed, fish are introduced to an area or escape, or when migration patterns change due to environmental conditions. The present results strengthen the observation regarding the need for conservation of this fish species. Further, the present study would serve as a baseline for future monitoring of the populations of *E. suratensis* in Kerala. Additional sampling of these populations in the future would allow the scientists to track any changes in the level of genetic variation in natural populations.

**Acknowledgement**

First author is indebted to the University Grants Commission (UGC), Government of India, New Delhi for the research fellowship to carry out the work. We thank the Director, Rajiv Gandhi Centre for Biotechnology, Thiruvananthapuram for the facilities received for the study.

**References**


