Evaluation of genetic pollution and genetic diversity in natural fisheries and aquaculture through molecular tools

Mohd Danish, Pooja Badoni, Karma Rinya, Supravat Maity, Anjan Chowdhury and Tushar Subhra Khanra

Abstract
Fish as a group, apart from its nutritional value from a biodiversity point of view, has the highest species diversity among all vertebrate taxa. Fishes exhibit enormous diversity in size, shape, biology and in the habitats they occupy. Fishes form a highly successful group of animals comprising more than 30,700 species inhabiting all seas, rivers, lakes, canals, dams, brackish water, estuaries and all places wherever is water. Major carp species contribute very significantly in total fish production of our country. Labeo rohita and and Cyprinus carpio var. communis are one of the most important culture species in aquaculture system and facing the problem of inbreeding in captive stocks and problem of overexploitation in wild stocks resulting reduction in level of heterozygosity and genetic diversity thereby it is important to analyze genetic structure of these species in order to devise and aid for their stock management and conservation. Unintentional inbreeding is a common default practice in hatcheries. A genetic marker is a site of heterozygosity for some type of silent DNA variation not associated with any measurable phenotypic variation. Such a “DNA locus,” when heterozygous, can be used in mapping analysis just as a conventional heterozygous allele pair can be used. Genetic variation in a species enhances the capability of organism to adapt to changing environment and is necessary for survival of the species. Various molecular provide different scientific observations which have importance in aquaculture practice recently such as: 1) Species Identification 2) Genetic variation and population structure study in natural populations 3) Comparison between wild and hatchery populations 4) Assessment of demographic bottleneck in natural population 5) Propagation assisted rehabilitation programmes.

Keywords: Genetic diversity, bottleneck and rehabilitation

Introduction
Almost 25% of global vertebrate diversity is accounted for by fish and India is the home for more than 11.72% of global fish biodiversity. Economically, fishes constitute a very important group of animals as a rich source of protein, liver oil and omega fatty acids. Fisheries sector plays a very important role in social and economic development by providing employment and nutritional security for the greater part of population of the country. Indian fisheries constitute about 5.17% of agriculture GDP and 0.9 % of the net GDP with total fish production of 10.79 million metric tonnes of which of 3.58 million metric tonnes from marine fisheries and 7.21 million metric tonnes from inland sector, out of which 6.489 million metric tonnes is from aquaculture(Ayyappan, et al., 2013) [7]. Owing to anthropogenic stresses, the fish availability from natural sources has been alarmingly declining world over and affecting sustainability of fisheries resources since their gene pools and genetic diversity is being eroded. The natural population of many fish species have experienced drastic reduction in number, largely due to the effects of overexploitation, habitat alterations including physiographic, abiotic and biotic features, escape of fish from fish farms and introduction of exotic species. This has adversely affected sustainability of many fisheries resources by eroding their gene pools and as a result genetic diversity. Genetic variation is an important feature of population both for short term fitness of individuals and long term survival of the population through allowing adaptation to changing environmental conditions (Ferguson et al., 1995) [29]. Determining the genetic structure is essential for developing controlled propagation, stock improvement and conservation plans. In terms of genetic management perspectives, the aim of natural fisheries management should be to conserve intra-specific genetic diversity for which description of the
genetic diversity of the concerned species is a pre-requisite for understanding the status and management requirements of the fish genetic resources. The genetic diversity of a species cannot be estimated from phenotypic data directly collected in the wild because of the possible occurrence of environmental effects which preclude accurate interpretation of observed variations. Now, genetic variations can be directly assessed through genetically controlled markers. Moreover, molecular genetic markers are powerful tools to detect genetic uniqueness of individuals, populations or species. These markers have revolutionized the analytical power necessary to explore the genetic diversity (Lakra et al., 2007) [42]. A molecular marker is a site of heterozygosity for some type of silent DNA variation not associated with any measurable phenotypic variation. Such a “DNA locus,” when heterozygous, can be used in mapping analysis just as a conventional heterozygous allele pair can be used. Several molecular tools have been used to assess genetic variation, determine population genetic structure and gene flow among fish species (Barroso et al., 2005) [12]. All organisms are subject to mutations because of normal cellular operations or interactions with the environment, leading to genetic variation (polymorphism). Genetic variation in a species enhances the capability of organism to adapt to changing environment and is necessary for survival of the species (Fisher, 1930) [52]. In conjunction with other evolutionary forces like selection and genetic drift, genetic variation arises between individuals leading to differentiation at the level of population, species and higher order taxonomic groups. Molecular genetic markers are powerful tools to detect genetic uniqueness of individuals, populations or species (Linda & Paul, 1995) [44]. These markers have revolutionized the analytical power, necessary to explore the genetic diversity (Hillis et al. 1996) [37]. Molecular markers can be classified into type I and type II markers. Type I markers are associated with genes of known function, while type II markers are associated with anonymous genomic regions (O’Brien, 1991) [57]. Under this classification, allozyme markers are type I markers because the protein they encode has known function. RAPD markers are type II markers because RAPD bands are amplified from anonymous genomic regions via the polymerase chain reaction (PCR). Microsatellite markers are also type II markers unless they are associated with genes of known function. In general, type II markers such as RAPDs, microsatellites, and AFLPs are considered non-coding and therefore selectively neutral. Such markers have found widespread use in population genetic studies to characterize genetic divergence within and among the populations or species (Brown and Epifanio, 2003) [13].

Protein markers
Analysis of allozyme loci remained one of the most popular approaches in examining population genetics and stock structure questions in fishes (Suneetha, 2000) [79]. The technique is rapid, relatively inexpensive and provides an independent estimate of level of variation within a population without an extensive morphological and quantitative survey (Menezes et al. 1993) [57]. Amino acid differences in the polypeptide chain of the different allelic forms of an enzyme reflect changes in the underlying DNA sequence. Depending on the nature of the amino acid changes, the resulting protein products may migrate at different rates (due to charge and size differences) when run through a gel subjected to an electrical field. Differences in the relative frequencies of alleles are used to quantify genetic variation and distinguish among genetic units at the levels of populations, species, and higher taxonomic designations. Disadvantages associated with allozymes include occasional heterozygote deficiencies due to null (enzymatically inactive) alleles and sensitive to the amount as well as quality of tissue samples. In addition, some changes in DNA sequence are masked at the protein level, reducing the level of detectable variation. Some changes in nucleotide sequence do not change the encoded polypeptide (silent substitutions) and some polypeptide changes do not alter the mobility of the protein in an electrophoretic gel (synonymous substitutions). At present 75 isozyme systems representing several hundred genetic loci are known (Murphy et al. 1996) [53].

Mitochondrial DNA markers
The development of DNA-based genetic markers has had a revolutionary impact on animal genetics. With DNA markers, it is theoretically possible to observe and exploit genetic variation in the entire genome. Popular genetic markers in the aquaculture community include allozymes, mitochondrial DNA, RFLP, RAPD, AFLP, microsatellite, SNP, and EST markers. The application of DNA markers has allowed rapid progress in aquaculture investigations of genetic variability and inbreeding, parentage assignments, species and strain identification, and the construction of high-resolution genetic linkage maps for aquaculture species. Well-designed studies using these genetic markers will undoubtedly accelerate identification of genes involved in quantitative trait loci (QTL) for marker-assisted selection.

Restriction fragment length polymorphism (RFLP)
RFLP markers (Botstein et al. 1980) [11] were regarded as the first shot in the genome revolution (Dodgson et al. 1997) [23], marking the start of an entirely different era in the biological sciences. Restriction endonucleases are bacterial enzymes that recognize specific 4, 5, 6, or 8 base pair (bp) nucleotide sequences and cut DNA wherever these sequences are encountered, so that changes in the DNA sequence due to deletions, base substitutions, or rearrangements involving the restriction sites can result in the gain, loss, or relocation of a restriction site. Digestion of DNA with restriction enzymes results in fragments whose number and size can vary among individuals, populations, and species. Most recent analyses replace the tedious Southern blot method with techniques based on the polymerase chain reaction (PCR). If flanking sequences are known for a locus, the segment containing the RFLP region is amplified via PCR. If the length polymorphism is caused by a relatively large (> approx. 100 bp depending on the size of the undigested PCR product) deletion or insertion, gel electrophoresis of the PCR products should reveal the size difference. However, if the length polymorphism is caused by base substitution at a restriction site, PCR products must be digested with a restriction enzyme to reveal the RFLP.

Random amplified polymorphic DNA (RAPD) markers
RAPD markers are the amplified products of less functional part of the genome that do not strongly respond to selection on the phenotypic level. Such DNA regions may accumulate more nucleotide mutations with potential to assess inter-population genetic differentiation (Mamuris et al. 2002) [49]. The amplification of genomic DNA by PCR with arbitrary nucleotide sequence primers, RAPD can detect high levels of
DNA polymorphisms (Williams et al., 1990) [91]. The technique detects coding as well as non-coding DNA sequences, and many of the most informative polymorphic sequences are those derived from repetitive (non-coding) DNA sequences in the genome (Haymer, 1994) [35]. Because 90% of the vertebrate nuclear genome is non-coding, it is presumed that most of the amplified loci will be selectively neutral. RAPD loci are inherited as Mendelian markers in a dominant fashion and scored as present/absent. RAPDs have all the advantages of a PCR-based marker, with the added benefit that primers are commercially available and do not require prior knowledge of the target DNA sequence or genome organization. Other advantages of RAPDs include the ease with which a large number of loci and individuals can be screened simultaneously. Shortcomings of this type of marker include the difficulty of demonstrating Mendelian inheritance of the loci and the inability to distinguish between homozygotes and heterozygotes. Analysis follows the assumption that populations under study follow Hardy-Weinberg expectations. In addition, the presence of paralogous PCR product (different DNA regions which have the same lengths and thus appear to be a single locus), low reproducibility due to the low annealing temperature used in the PCR amplification, have limited the application of this marker in fisheries science (Virgin and Waldman, 1994)[92].

Single nucleotide polymorphism (SNP)

Single nucleotide polymorphism (SNP) describes polymorphisms caused by point mutations that give rise to different alleles containing alternative bases at a given nucleotide position within a locus. SNPs are becoming a focal point in molecular marker development since they represent the most abundant polymorphism in any organism’s genome (coding and non-coding regions), adaptable to automation, and reveal hidden polymorphism not detected with other markers and methods (Liu and Cordes, 2004) [45]. Theoretically, a SNP within a locus can produce as many as two alleles, each containing one of two possible base pairs at the SNP site. Therefore, SNPs have been regarded as biallelic. SNP markers are inherited as co-dominant markers. Several approaches have been used for SNP discovery including SSCP analysis [Hecker et al. 1999][56], heteroduplex analysis, and direct DNA sequencing. DNA sequencing has been the most accurate and most used approach for SNP discovery.

Microsatellite markers

Microsatellites consist of multiple copies of tandemly arranged simple sequence repeats (SSRs) that range in size from 1 to 6 base pairs [e.g., ACA or GATA; Hecker, 1999] [36]. Microsatellites tend to be arranged simple sequence repeats (SSRs) that range in size from 1 to 6 base pairs [e.g., ACA or GATA; Hecker, 1999] [36]. Microsatellites are easily distributed in the genome on all chromosomes and all regions of the chromosome. However, data from whole genome sequencing has somewhat contradicted this statement. They have been found inside gene coding regions (Liu et al. 2001[55]), introns, and in the non-gene sequences. Most microsatellite loci are relatively small, ranging from a few to a few hundred repeats. Regardless of specific mechanisms, changes in numbers of repeat units can result in a large number of alleles at each microsatellite locus in a population. Microsatellites have been inherited in a Mendelian fashion as codominant markers. Microsatellites were found to be informative in several species, which showed almost no variation at other markers (Taylor et al. 1994) [33]. However, use of microsatellite markers involves a large amount of up-front investment and effort. Each microsatellite locus has to be identified and its flanking region sequenced to design of PCR primers. Due to polymerase slippage during replication, small size differences between alleles of a given microsatellite locus (as little as 2 bp in a locus comprised of di-nucleotide repeats) are possible. Microsatellites recently have become an extremely popular marker type in a wide variety of genetic investigations.

Expressed sequence tags (ESTs)

Expressed sequence tags (ESTs) are single-pass sequences generated from random sequencing of cDNA clones (Adam et al. 1991) [3]. The EST is use to identify genes and analyze their expression by means of expression profiling. It helps for rapid and valuable analysis of genes expressed in specific tissue types, under specific physiological conditions, or during specific developmental stages. ESTs offer the development of cDNA microarrays that allow analysis of differently expressed genes to be determined in a systematic way (Wang et al. 1999) [86]. For genome mapping, ESTs are most useful for linkage mapping and physical mapping in animal genomics. In spite of its popularity in mammalian genome mapping (Korwin-Kossakowska, 2002 [41]) radiation hybrid panels are not yet available for any aquaculture species. Development of radiation hybrid panels from aquaculture species is not expected in the near future, given the fact that physical mapping using BAC libraries can provide even higher resolution and the fact that BAC libraries are already available from several aquaculture species. Therefore, ESTs are useful for mapping in aquaculture species only if polymorphic ESTs are identified (Liu, 1999) [46].

Application of molecular markers in species identification

The inter-specific genetic divergence established through species specific diagnostic molecular markers provides precise knowledge on phylogenetic relationships Backer, 2002[2]; Asensio et al. 2002 [6] and also resolve taxonomic ambiguities (Rocha-Olivares et al. 2000[70]; and Rasmussen et al. 2003) [67]. These markers can be used to detect hybrid and introgressed or backcrossed individuals, distinguish early life history stage of morphologically close species (Olivar et al. 1999) [59] both in hatchery and in natural populations. Species-specific allozyme markers have been identified in many fishes [Tilapia: 72; Sciaenid: 73; Anguilla sp: 79; Mugilidae: 80] Specific diagnostic allozyme loci were used for different species: apache trout (Oncorhynchus apache), cutthroat (Oncorhynchus clarki) and rainbow trout (Oncorhynchus mykiss) (Carmichael et al. 1993) [16] and Gambusia affinis and G. holbrooki (Wooten and Lydeard, 1990) [93]. Allozyme markers have also been used for individual classification in cyprinid species Zacco pachycephalus and Z. platypus (Wang-Hung et al. 1997) [87], in cyprinodontid species V. letourneuxi and V. hispanica (Perdices et al. 1996) [64], in mullets Mullus barbatus and M. surmuletus (Mamuris et al. 1998) [49] and hare species Merluccius australis and M. hubbsi (Roldán and Pla, 2001) [72]. Species-specific diagnostic RAPD fingerprints were generated in several fish species and their taxonomic relationship has been analyzed. RAPD markers were characterized to identify five species of family Cyprinidae: Chondrostoma lemmingii, Leuciscus pyrenaicus, Barbus bocagei, Barbus comizo, all endemic in the Iberian Peninsula and introduced Alburnus alburnus (Callejas and Ochando, 2002) [15], for studying genetic relationship and
diversities in four species of Indian Major carps (family Cyprinidae): rohu (Labeo rohita), kalbasu (L. calbasu), catla (Catla catla) and mrigal (Cirrhinus mrigala) (Barman et al. 2003) [9].

**Genetic variation and population structure study in natural populations**

Geographic distance and physical barriers enhance reproductive isolation by limiting the migration and increase genetic differentiation between populations (Ryman, 2002) [74]. Impact of migration and gene flow on genetic differentiation also depends upon effective size of receiving population and number of migrants. Increased computational power and mathematical models have enhanced the scope of conclusions that can be drawn out of genotype data generated through molecular markers. Some of the possibilities are assignment of migrants (Piry et al. 2004) [65], determination of genetic bottlenecks (Luikart and Cornuet, 1998) [45], effective breeding population estimates (Luikart and Cornuet, 1999) [48]. Population genetic structure has been investigated using allozyme markers in many fish species, Oncorhynchus gorbuscha (Efremov, 2002) [27] and Tenualosa ilisha (Salini et al. 2004) [79]. Fifteen random primers were used to analyze the genome DNA of Jian carp (Cyprinus carpio var jian) by the RAPD technique (Dong et al. 2002) [24]. Study on cold tolerant traits for common carp Cyprinus carpio was conducted by Chang et al. (2003) [18] and nine RAPD-PCR markers associated with cold tolerance of common carp were identified. Population structure has been examined using microsatellite markers of sockeye salmon (Nelson et al. 2003) [54], Chinook salmon (Beacham et al. 2003) [10] and Arctic charr populations (Brunner et al. 1998) [14]. Genetic variation have been assessed using microsatellite genetic markers to identify the population structure of brook charr, Salvelinus fontinalis (Adams and Hutchings, 2003) [1] and 14 populations of northern pike (Esox lucius) in the North Central United States and in six populations from Quebec, Alaska, Siberia, and Finland (Senanan and Kapuscinski, 2000) [78]. Based on five microsatellite loci, the genetic structure of endangered fish species Anaecypris hispanica was studied in eight distinct populations in the Portuguese Guadiana drainageto determine levels of genetic variation within and among populations and suggested implications for conservation of the species (Salgueiro et al. 2003) [74]. Combination of allozyme and microsatellites was used to investigate genetic divergence in Salmo trutta (Palm et al. 2003) [61] and Salmo salar (Elliott and Reilly, 2003) [28]. Alarcon et al. (2004) [3] represents population genetic analysis of gilthead sea bream (Sparus aurata), Kanda and Allendorf (2001) [39] examine population genetic structure of bull trout Salvelinus confluentus using a combination of allozyme, microsatellite and mtDNA variation.

**Comparison of genetic variation between wild and hatchery populations**

Molecular markers also find application in aquaculture to assess loss of genetic variation in hatcheries through, comparison of variation estimates between hatchery stocks and wild counterparts. The information is useful obtained in monitoring farmed stocks against inbreeding loss and to plan genetic upgradation programmes. A major aspect such studies address is concerned with the assessment of farm escapes into the natural population and introgression of wild genome. All wild-unstocked samples were highly differentiated populations and significantly different from each other and from hatchery samples. Genetic diversity was investigated using microsatellites between farmed and wild populations of Atlantic salmon (Norris et al. 1999) [59]. Farmed salmon showed less genetic variability than natural source population in terms of allelic diversity. Variation in allozymes and three microsatellite loci was assessed in populations of wild and cultured stocks of Sparus aurata (Palma et al. 2004) [60] and Sparus auratus (Alarcon et al. 2004) [3]. The microsatellite heterozygosity value were high in wild, but lower in the cultured samples.

**Application of microsatellites in population structure analysis in fisheries and aquaculture**

Highly polymorphic microsatellite markers have great potential utility as genetic tags for use in aquaculture and fisheries biology. They are powerful DNA markers for quantifying genetic variations within and between populations of species (Weber, 1990) [88]. The qualities of microsatellites make them very useful as genetic markers for studies of population differentiation and stock identification (Liu and Cordes, 2004) [45], in kinship and parentage exclusion (Webster and Reichart, 2005) [89] and Hansen et al. 2001) [74] and in genome mapping (Sanetra et al. 2009) [77]. Microsatellites are also being used as genetic markers for identification of population structure, genome mapping, pedigree analysis, and to resolve taxonomic ambiguities in many other animals besides fishes (Nikbakht et al. 2013 [15]; Arias-Perez et al. 2012 [3]; Fernandes et al. 2012 [31]; Upadhyay et al. 2012 [85]; Joshi et al. 2012 [38]; Xu and Liu 2011 [94], Supungul et al. 2006) [80]. The broad areas of applications of microsatellite markers are depicted in the development of polymorphic microsatellite markers to determine the population structure of the Patagonian toothfish, Dissostichus eleginoides, has been reported by (Rogers et al. 2006) [73]. Similarly, Appleyard et al. (2002) [4] examined seven microsatellite loci in the same species of Patagonian toothfish from three locations in the Southern Ocean. Recently, Larsen et al. (2011) [43] showed differences in salinity tolerance and its gene expression in two populations of Atlantic cod (Gadus morhua). Drinan et al. (2011) [25] reported 20 microsatellites for determining the patterns of population genetic variation in westslope cutthroat trout, Oncorhynchus clarkia lewissii in 25 populations from four rivers. Davies et al. (2011) [20] identified 12 microsatellite loci in tuna species of genus Thunnus and investigated genetic polymorphism at these loci in North Atlantic and Mediterranean Sea populations. Similarly, several authors reported population genetic structure of different species of catfish; few of them are in the farmed catfish from Tamaulipas, Mexico (Peralas-Flores et al. 2007) [63], in neotropical catfish (Riboli et al. 2012) [69]; in Pseudoplatystoma reticulatum (De Abreu et al. 2009) [21], O’Connell et al. (1997) [18] reported the investigation of five highly variable microsatellite loci for population structure in Pacific herring, Clupea pallasi, collected from 6 sites in Kodiak Island. Similarly, many others have reported studies of polymorphic microsatellite loci to evaluate population structure of different fish species. Thus microsatellite markers have wide range of applications in population genetics and fisheries management. Salzburger et al. (2002) [76] reported a case of introgressive hybridization between an ancient and genetically distinct cichlid species in Lake Tanganyika that led to the recognition of a new species. DeWoody and Avise
reported microsatellite variation in marine, fresh water, and anadromous fishes compared with other animals. Gopalakrishnan et al. (2009) carried out characterization of dinucleotide microsatellite repeats in *Laboe rohitua*. As these factors would lead to a reduction in reproductive fitness (Padhi and Mandal, 2000), efforts to increase the genetic diversity of the fish species should be given high priority for conservation of the species, based on genetic principles as mentioned below.

1. The effective population size (Ne) should be maintained as large as possible to maximize the contribution of a large number of adults for reproduction so as to maintain natural genetic variability.

2. The causative factors that reduce the effective population size such as overexploitation should be controlled at the earliest.

3. No artificial gene flow between distinct stocks should be created by means of haphazard stocking and rehabilitation programs.

4. The rehabilitation strategy should also include means (screening the population, using genetic markers) to monitor impact of such program.

5. The natural populations of the endangered species can be enhanced by “supportive breeding.” In this program, a fraction of the wild parents are bred in captivity and the progeny are released in natural waters.

6. Brood stock of fish species collected from different rivers must be tagged and maintained in separate ponds in the holding facility.

7. Effective breeding population size and sex ratio should not be restricted. To achieve this, collection of different size/year classes at different time intervals is to be preferred over the same size/year class.

8. Use of cryopreserved milt, collected from different males and pooled, would be useful for increasing the effective population size and recovery of endangered populations of fish species. In comparison to the captive breeding program, the gene banking through sperm cryopreservation is relatively cheaper, easy to maintain, and less prone to risk due to system failure or mortality due to diseases. Therefore, it should serve as a useful adjunct to the captive breeding program.

9. Different genetic stocks should be bred separately and ranched in the same rivers from where they are collected.

10. Stretches of rivers harboring resident population or that can serve as a potential sanctuary, may be selected for ranching of fish populations.

11. Assessing the impact of ranching through monitoring the parameters like catch per unit effort/area through experimental fishing should be done.

References


19. Danzmann RG, Ihssen PE. “Aphylogeographic survey of


35. Haymer DS. Random amplified polymorphic DNAs and microsatellites: What are they, and can they tell us anything we don’t already know? Annals of Entomological Society of American 1994;87:717-722.


51. Menezes MR, Naik S, Martins M. Genetic
63. Perales-Flores LE, Sifuentes-Rincón AM, García de León FJ. “Microsatellite variability analysis in farmed catfish (Ictalurus punctatus) from Tamaulipas, Mexico,” Genetics and Molecular Biology 2007;30(3):570-574.
79. Suneetha BK. Interspecific and inter specific genetic variation in selected mesopelagic fishes with emphasis on microgeographic variation and species characterization. Dr. Scient. Dissertation, Department of Fisheries and Marine Biology, University of Bergen, Bergen, Norway 2000.


