Chapter
16

Applications of Molecular Markers in Fisheries and Aquaculture

Sandhya Sukumaran and A. Gopalakrishnan

Marine Biotechnology Division
ICAR -Central Marine Fisheries Research Institute, Kochi

Organisms are characterized by unique biological attributes which enhance their fitness and survival to a particular environment. The driving force for enhanced survival and fitness is the genetic variation inherent in an individual as well as in a population. The information regarding genetic diversity and variation has wide application in research on evolution, conservation and management of natural populations. The advent of DNA cloning and sequencing methods have contributed immensely to the development of molecular taxonomy and population genetics over the last 2 decades. These modern methods have revolutionized the field of molecular taxonomy and population genetics with improved analytical power and precision.

Molecular markers can be characterized as Type I and Type II markers; Type I markers are associated with genes of known function and type II markers are associated with genes of unknown function. Allozyme markers are type I markers as the proteins they encode are associated with some functions. Microsatellites and other neutral markers are type II markers unless they are associated with genes of some known function.

Allozyme

Allozyme electrophoresis is a method which can identify genetic variation at the level of enzymes that are directly encoded by DNA. Protein variants called allozymes originates from

allelic variants and they will differ slightly in electric charge. Allozymes are codominant markers having been expressed in a heterozygous individual in a Mendelian way. Thus allozyme analysis provides us with data on single locus genetic variation which can answer many questions about fish and fish populations. To detect allozyme variation, the first step is to extract allozymes from tissues using specific protocols. Then the variation is detected through electrophoresis in an acrylamide or cellulose acetate gel. Individuals that are homozygous show a single band whereas heterozygous individuals show two bands. Allozymes are one of the most studied form of molecular variation due to their simplicity, low cost and the requirement of little specialized equipment. Any kind of soluble protein is suitable for allozyme analysis. A large number of loci can be screened at a time. The limitations of this technique include requirement of a large amount of tissue and consequently this method could not be applied when the organisms are small (for e.g.; larval forms). The tissue sampling method is invasive and so the fish needs to be sacrificed and the tissue needs to be stored cryogenically. A point mutation in a nucleotide sequence may not result in a change in amino acid at all and thus could not be detected by protein electrophoresis. In addition to that, a change in DNA that results in a change in amino acid will not result in the overall charge of the protein and therefore is not detected. In spite of their limitations, the use of allozyme analysis has been widespread in fisheries mainly in fish systematics, population structure, conservation genetics, mixed stock fishery analysis and forensic analysis.

Mitochondrial DNA markers

Mitochondrial DNA is non- nuclear DNA in the cell having located in within organelles in the cytoplasm called mitochondria. Mitochondrial DNA is maternally inherited with a haploid genome. The entire genome undergoes transcription as one single unit. They are not subjected to recombination and thus they are homologous markers. They are selectively neutral occurring in multiple copies in each cell. Mitochondrial DNA is physically separate from the rest of cell's DNA and so it is relatively easy to isolate from any tissue or blood sample. Due to the maternal inheritance of mitochondrial DNA, the effective population size is smaller than nuclear DNA and so mitochondrial DNA variation is more sensitive to population bottlenecks and hybridizations. The differences in the nucleotide sequence of DNA molecule in the mitochondria can be

determined directly or indirectly by several methods. Many population genetic studies have employed RFLP (Restriction Fragment Length Polymorphism) analysis of mitochondrial DNA for understanding population genetic variation either by digesting the whole purified mtDNA with restriction endonucleases or by DNA sequencing of small segments of mtDNA molecule obtained by PCR amplification. These techniques with increased resolution and maximum information have made mtDNA analysis very popular.

The newly emerged sequencing technologies have enabled direct sequencing of mitochondrial genes and several sets of universal primers have been developed from conserved sequence regions. Slow evolving gene regions are being used for inter species comparisons and fast evolving gene regions for population comparisons. The only non-coding region of mtDNA is D-loop region and this region is fast evolving and mostly used for population comparisons. The cytochrome b and ND-1 and ND-5/6 gene regions are also being used widely. Mitochondrial Cytochrome C Oxidase I gene has been identified as the universal barcode for species level identification due to its conserved nature across a wide range of taxa. DNA barcodes are segments of approximately 600 base pairs of the mitochondrial COI gene which is a fast, efficient and inexpensive technique helpful in cataloguing the biodiversity. During the last two decades, mitochondrial DNA genes have found widespread application fish taxonomy, biology and population genetics.

Arbitrary Nuclear DNA markers

Arbitrary markers are used when we target a segment of DNA of unknown function. The widely used methods of amplifying unknown regions are RAPD (Random Amplified Polymorphic DNA) and AFLP (Amplified Fragment Length Polymorphism) DNA. RAPD uses an arbitrary primer which can amplify anonymous loci. It is fast, cheap and shows very high amount of polymorphism and this marker does not require knowledge of the genetic makeup of the organism. The major drawback with RAPD markers is the lack of reproducibility and repeatability and the large number of products generated. RAPD is a dominant marker and so homozygous and heterozygous states cannot be differentiated and these patterns are sensitive to slight changes in amplification conditions. Amplified Fragment Length Polymorphism (AFLP) markers

combine the benefits of both RFLP and RAPD. The total genomic DNA is digested using two restriction enzymes. Double–stranded nucleotide adapters are ligated to the ends of DNA fragments to serve as primer binding sites for PCR amplification. Primers complementary to the adapter and restriction site sequence, with additional nucleotides at the 3'-end, are used as selective agents to amplify a subset of ligated fragments. The presence of absence of DNA fragments are detected on polyacrylamide gels and thus polymorphisms are studied.

Specific Nuclear DNA markers

Variable Number of Tandem Repeat is a segment of DNA that is repeated tens or even hundreds to thousands of times in nuclear genome of eukaryotes. They repeat in tandem; vary in number in different loci and differently in individuals. There are two main classes of repetitive and highly polymorphic DNA; minisatellite DNA referring to genetic loci with repeats of length 9-65 bp and microsatellite DNA with repeats of 2-8 bp (1-6) long. Microsatellites are much more numerous in the genome of vertebrates than mini satellites. They are widely used in population genetics of fishes and aquatic invertebrates. Minisatellites can be classified into multilocus and single-locus minisatellites. Multilocus minisatellites are composed of tandem repeats of 9-65 base pair and have a total length ranging from 0.1 to 7kb. Minisatellite loci are used mainly in parentage analysis. They are less useful for population genetic analysis unless we use large sample sizes. The complexity of mutation processes undergone by minisatellite loci is also a limitation. Due to the difficulties in the interpretation of multilocus fingerprints, the research work were concentrated on single locus minisatellite probes and this procedure required reasonable quantities of high-quality DNA. These single locus minisatellite probes have been very useful and successful in detecting genetic variations within and between populations. It has also been used in fisheries for forensics, parentage, genetic identity, estimating mating success and confirming gynogenesis.

Microsatellites

A microsatellite is a simple DNA sequence which is repeated several times across various points in the DNA of an organism. These repeats are highly variable and these loci can be used as markers. Microsatellite occur once in every 10 kbp while minisatellite loci occur once in every

1500 kbp in fishes and due this, microsatellites are more useful in genome mapping and population genetics studies. They are highly variable, non-coding and selectively neutral and the basic assumption while using microsatellite loci is that the predicted amount of sequence divergence between units of interest is directly related to length of time since separation. Microsatellites are codominant markers which are inherited in a Mendelian fashion and they are highly evolving with 10^{-3}-10^{-4} mutation/generation. The high levels of polymorphism shown by microsatellites have made them one of the most popular genetic markers. Cross amplification with primers developed in closely related species is also possible which minimizes the cost associated with detecting microsatellite sequences in a different species. The analysis of microsatellite loci involves DNA extraction, amplification of the microsatellite loci using specific primers in a PCR machine and examination of the bands using poly acrylamide gel electrophoresis. The recent introduction of automated genotyping machines has made the analysis of size polymorphisms of microsatellite loci with automated genotyping using labeled primers. The use of large number of samples and loci is now possible due to automated genotyping which has increased precision and speed with microsatellite analysis. The constraints of using microsatellite markers are the presence of null alleles and presence of stutter bands. Null alleles are found when mutations occur at primer binding sites of microsatellite locus. The presence of null alleles reduce accuracy especially in parentage or relatedness analysis and assignment tests and the best option is to discard loci showing null alleles. Stutter bands occur when a ladder of bands differing between 1-2 bp is seen and these occur due to slipped strands impairing during PCR or incomplete denaturation of amplification products. Tri-nucleotide and tetra nucleotide repeats usually do not show significant amounts of stuttering. Microsatellite markers are used in fisheries and aquaculture for phylogenetics and phylogeography, population genetic structure, biodiversity conservation, stocking impacts and hybridization. It is also being increasingly used for forensic identification of individuals, genome mapping and determination of kinship and behavioral patterns.

Single Nucleotide Polymorphisms

Single nucleotide polymorphisms arise due to single nucleotide substitutions (transitions/transversions) or single nucleotide insertions/deletions. These point mutations give

rise to different alleles with alternative bases at a particular nucleotide position. SNP,s are the most abundant polymorphisms in the genome (coding and non-coding) of any organism. These single nucleotide variants can be detected using PCR, microchip arrays or fluorescence technology. They are considered as next generation markers in fisheries and can be employed for population genetics studies, genomics studies and for detection of diseases.

DNA microarrays or DNA chips

DNA microarray consists of small glass microscope slides, silicon chip or nylon membranes with many immobilized DNA fragments arranged in a standard pattern. A DNA microarray can be utilized as a medium for matching a reporter probe of known sequence against the DNA isolated from the target sample which is of unknown origin. Species-specific DNA sequences could be incorporated to a DNA microarray and this could be used for identification purposes. DNA extracted from a target sample should be labeled with a specific fluorescent molecule and hybridized to the microarray DNA. When the hybridization is positive a fluorescent signal is detected with appropriate fluorescence scanning/imaging equipment.

Expressed Sequence Tags (ESTs)

ESTs are single-pass sequences which were generated from random sequencing of cDNA clones. ESTs can be used to identify genes and analyze their expression by means of expression analysis. Fast and reliable analysis can be made for the genes expressed in particular tissue types under specific physiological conditions or developmental stages. Differentially expressed genes could be identified using cDNA microarrays in a systematic way. ESTs are most valuable for linkage mapping.

Applications of molecular markers in fisheries

Inter specific and intra specific variations

Molecular genetic markers can be used as a supplementary marker system which will increase resolution in taxonomic research. The molecular evolution among taxa is highly variable and the extent of divergence in DNA or genes can be taken as the basis for differentiation among species. The morphological and ecological characters may diverge at a faster rate compared to

genetic differentiation at neutral loci and so sometimes we may observe poor correlation between morphological traits and gene divergence. Thus molecular markers have to be interpreted critically among different sets of traits. Molecular markers are most useful when some species which occur in mixed catches is to be identified and where morphological identification is very difficult. Processed fish products (filleted, smoked or salted) and early life stages of fishes like planktonic eggs and larval forms cannot be identified using morphological characters and molecular markers can be employed in such instances. Endangered and threatened whales, sharks and dolphins which are dead and stranded can also be identified using these methods as in most cases morphological identification is not possible. Within a species, genetic differences are more than between populations and so identification is possible even when sample size is small (3-5). Molecular markers have also been used in sub-species identification. Some of the high evolving loci will show more divergence within species and so these loci can be used for finding out intra-specific variations.

Phylogenetic and Phylogeograhical studies

Phylogenetic studies assess the historical processes which affect relationships and phylogeographic studies assess the geographical distributions. Phylogenetic and phylogeographic studies started with the introduction of mtDNA markers in population genetic analyses. The evolutionary history of groups of fishes could be reconstructed which will give vital information regarding historical demography. Information regarding conservation units and ecological patterns could also be derived using phylogenetic studies. Mitochondrial DNA analysis has been used widely as a powerful tool for intraspecific phylogenetic patterns inference in many animal species. The high levels of mutation rate, smaller levels of effective population size and predominantly maternal inheritance of mtDNA will provide greater power to identify population structure. The lack of recombination and low efficiency of repair mechanisms induces high rate of evolution in mtDNA which makes this molecule highly useful in phylogenetic analyses. MtDNA has been used to resolve relationships among species that had diverged as much as 8-10 million years.

Structure of populations: between and within populations

Identification of stock structure is a very important issue in fisheries management and conservation programmes. Stocks are groups within species which are reproductively isolated with different physiological and behavioural patterns. Morphological and meristic features have genetic and environmentally influenced components and so morphological and meristic data should be used in conjunction with genetic data. Thus when a combination of these methods is used, we get information regarding actual genetic differences and important environmental effects on phenotype. Microsatellite DNA is the most favored DNA marker for stock structure studies due to their high rate of evolution. The non-coding region of mtDNA (D-loop) has also been used for stock structure studies due to their high rate of evolution. Mixed fisheries are comprised of subunits of different populations, different life stages or individuals from different stocks. Data from microsatellite markers, allozymes or mitochondrial DNA markers is useful for identifying the origin of stock components from mixed stock fisheries.

Genetic tagging/marking

There are instances when we need to mark individual fish for various purposes like tracking movement or migration, estimating population size or evaluating contributions of individual stock to mixed fishery. Physical tags are not heritable and so cannot be employed for generations. Genetic marking by finding a rare allele in an individual or populations and following them over generations will provide information regarding the contribution of hatchery programme on harvest, proportion of stocked individuals on growth of targeted population and identifying migrants from source populations.

Forensic investigation

In samples where morphological identification is not possible like dead or stranded fishes, preserved or canned fish flesh and fish fillets, molecular markers could be employed for identification and certification. Forensics uses scientific methods to draw inferences about past events and it is being increasingly used in certification of fishery products and detection of illegal trading in fish and fisheries products. Deliberate or accidental release into natural waters also could be monitored using molecular tools.

Studying the trophic relationships

The most essential component of any ecological study is determining trophic relationships within an ecosystem and data on diet composition is very vital. Identifying the diet components of species level is very difficult as most the morphological features might get lost due to partial or complete digestion. In many cases solid remains do not exist inside the gut and this makes identification difficult. Molecular methods could be used for diet analysis studies as it is possible to extract DNA from partially digested samples.

Analysis of ancient DNA

Retrieval of DNA sequence information from preserved samples in museums, fossil remains, archaeological finds and other unusual sources of DNA is now possible with several methods. These types of studies have improved our understanding regarding the evolutionary relationships among different taxa.

Applications in aquaculture

Molecular markers have wide range of applications in aquaculture mainly; in genetic identification and discrimination of hatchery stocks, finding out inbreeding events, assignment of progeny to parents using genetic tags, finding out quantitative trait loci, marker assisted selection for selective breeding trials and assessment of the effect of polyploidy induction and gynogenesis. Genetic variability can also be assessed between and within stocks using molecular markers. They are also useful in determining the contribution of possible parents in mass spawning events. Genome mapping and identification of quantitative trait loci (QTL), (the locations of commercially important quantitative traits) is also possible. QTLs are important in breeding programmes and they are detected by analyzing phenotypes with linked marker maps and identification of markers related to QTLs can provide information on relatedness between strains and families and this knowledge is useful in marker-assisted programmes to improve production related traits. Disease diagnosis is another major area where the benefits of molecular markers can be successfully harnessed. PCR assays for different pathogens have become inexpensive, safe and user friendly in many diagnostic laboratories. Several PCR based disease diagnosis methods have been developed for pathogens like white spot syndrome virus

(WSSV), channel catfish virus (CCV), infectious hematopoietic necrosis virus (IHNV), infectious pancreatic necrosis virus (IPNV), viral hemorrhagic septicemia virus (VHSV), viral nervous necrosis virus (VNNV) and several other diseases. Genes that are resistant to pathogens like Major Histo Compatibility genes (MHC) can be identified and used for selection programmes to produce disease resistant strains of fin or shell fishes.

Conclusion

Molecular markers find wide range of application fisheries and aquaculture and their application has revolutionized the field of fish genetics. The choice of a marker type should be made cautiously in each case so as to maximize the quality of output. No single molecular marker is superior to any other and a combination of markers is always suitable. Of late, markers developed and screened using next generation sequencing technologies are increasingly being used in fish genetics. There is an increasing global demand for aquaculture products and modern molecular methods and molecular genetics could play a major role in bringing out quality and sustainability to aquaculture.

Suggested reading

Askari, G.H., Shabani, H. and Miandare, H.K. Application of molecular markers in fisheries and aquaculture. *Scientific Journal of Animal Science*, 2(4): 82-88.

Chauhan, T and Rajiv. K. 2010. Molecular markers and their application in fisheries and aquaculture. *Advances in Bioscience and Biotechnology* 1: 281-291.

Davis, G.P. and Hetzel, D.J.S. 2000. Integrating molecular genetic technology with traditional approaches for genetic improvement in aquaculture species. *Aquaculture Research*, 31: 3-10.

Hallerman, E.M. 2006. Use of molecular tools for research and improvement of aquaculture stocks. *The Israeli Journal of Aquaculture – Bamidgeh*, 58(4): 286-296.

Liu, Z.J. and Cordes, J.F. 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* 1-37.

Okumus, I. and Ciftci, Y. 2003. Fish population genetics and molecular markers: II- Molecular markers and their applications in fisheries and aquaculture. *Turkish Journal of Fisheries and Aquatic Sciences* 3: 51-79.

Presti, R.L., Lisa,C. and Stasio, L.D. 2009. Molecular genetics in aquaculture. *Italian Journal of Animal Science* 8: 299-313.

Teletchea, F. 2009. Molecular identification methods of fish species: reassessment and possible applications. *Reviews in Fish Biology and Fisheries* 19: 265- 293.