Molecular Tools and Techniques in Marine Fish Identification

Sandhya Sukumaran

ICAR-Central Marine Fisheries Research Kochi

Email: sukumaransandhya@yahoo.com

The well-being and persistence of various organisms in a specific environment rely on the distinctive biological characteristics of each species. The intrinsic genetic diversity within species and populations serves as a crucial factor for thriving amidst dynamic environmental conditions. Recognizing the genetic and morphological variations is essential for ensuring the sustainability of ocean resources, discerning adaptive evolution patterns, and effectively managing these resources. The accurate identification of fish species is paramount for sustainable fisheries management, preserving biological complexity. Determining stock structure is vital for formulating targeted management strategies. Recent advancements in molecular tools have facilitated precise identification and delineation of stock structure, empowering conservationists to plan and implement effective management and conservation measures.

Molecular taxonomic tools

Over the past decade, significant progress has been made in molecular taxonomic research, thanks to the emergence of advanced sequencing methods. The capabilities and accuracy inherent in molecular taxonomic investigations have greatly enhanced our understanding across various realms of biological science.

Various molecular markers are currently employed to decipher patterns of diversity within and between marine populations, categorized as Type I and Type II. A marker falls into the Type I category when linked to genes of known function, and into the Type II category when associated with genes of unknown function

Type I markers

Allozyme markers are type I markers as they are associated with genes of known functions. DNA encodes allozymes and genetic variation at the level of enzymes can be detected using allozyme electrophoresis. Allozymes are protein variants which originate from allelic variants which differ in electric charge and these variations could be detected using electrophoresis. Allozymes are codominant markers expressed in

heterozygous individuals in a Mendelian fashion. Information regarding single locus genetic variation can be gathered using allozyme analysis which can answer many questions regarding intra- and inter specific diversity of fish populations. To detect variations in allozyme pattern, allozymes have to extracted from tissues following standard protocols and variations detected through electrophoresis in an acrylamide or cellulose acetate gel. A single band will be present if individuals are homozygous and double bands when individuals are heterozygous. Allozymes have been extensively studied and used for many investigations due to their simplicity and cost effectiveness as any kind of soluble protein can be used for allozyme analysis. Many numbers of loci can be screened at a time using allozyme markers. Major limitations with the use of allozyme analysis include requirement of large amount of tissue which impedes its use with smaller organisms like larval forms. The tissue sampling method is invasive and hence every time the fish has to be sacrificed and tissue stored cryogenically. Point mutations in nucleotide sequences could not be detected using protein electrophoresis as such mutations may not result in change in the amino acid composition. In spite of all these limitations, allozymes have been widely used for fishery biology investigations like fish systematic, population genetic structure, conservation genetics and forensic applications.

Mitochondrial DNA markers

Mitochondrial DNA resides within the cytoplasm of organelles called mitochondria, distinguishing it as non-nuclear DNA. Functioning as a haploid genome, it undergoes maternal inheritance, and transcription occurs as a single unit. Mitochondrial DNA remains free from recombination events, is selectively neutral, and exists in multiple copies within each cell. Its isolation from various tissues or blood samples is straightforward due to its physical separation from the cell's nuclear DNA. The use of mitochondrial DNA facilitates the detection of population bottlenecks and hybridizations, given its smaller effective population size resulting from maternal inheritance.

Molecular taxonomic methods employing mitochondrial DNA often utilize Restriction Fragment Length Polymorphism (RFLP) based on length polymorphisms or sequence variations to reveal genetic diversity patterns. RFLP involves digesting mitochondrial DNA with restriction enzymes, visualizing resulting polymorphisms through electrophoresis, and has been widely applied in understanding species-specific patterns in marine fishes. Sequencing technologies have popularized the sequencing

of mitochondrial DNA regions, with universal primers targeting conserved sequences for broad applicability. Comparisons between species can be conducted using slowevolving gene regions, such as the universal barcode Cytochrome C Oxidase 1, while fast-evolving regions like the D-loop are employed for intra-specific or population comparisons. Other gene regions, including cytochrome b and ND-1 and ND-5/6, are also utilized for intra-specific comparisons. Mitochondrial DNA genes play a pivotal role in fish systematics and population genetics.

Nuclear DNA markers

Nuclear DNA markers can be categorized as arbitrary and specific depending on the gene regions to be amplified.

Arbitrary markers

Arbitrary markers, such as RAPD (Random Amplified Polymorphic DNA) and AFLP (Amplified Fragment Length Polymorphism), play a significant role in genetic analysis. RAPD utilizes an arbitrary marker to amplify genomic DNA regions, resulting in a high level of polymorphism. This approach is rapid, cost-effective, and efficient, requiring no prior knowledge of specific genes or the genome. However, RAPD markers exhibit limitations in terms of reproducibility and repeatability, generating multiple products simultaneously. These markers cannot distinguish between homozygous and heterozygous states, and slight variations in amplification conditions lead to changes in band patterns. AFLP markers combine Restriction Fragment Length Polymorphism (RFLP) and RAPD techniques. Genomic DNA is digested with two restriction enzymes, followed by the ligation of double-stranded nucleotide adapters to the DNA fragments' ends, serving as primer binding sites for PCR amplification. Selective agents, represented by primers complementary to the adapter and sequences at the restriction site with additional nucleotides at the 3′ end, amplify a subset of ligated fragments. Polymorphisms and the presence or absence of DNA fragments are then detected on polyacrylamide gels.

Specific markers

Variable Number of Tandem Repeats (VNTRs) refers to segments of DNA that are repeated numerous times within the nuclear genome of eukaryotes, occurring in tens, hundreds, or thousands of repetitions. These repeats are tandem and exhibit variation in number across different loci of the genome and among individuals. Repetitive DNA

can be categorized into minisatellite and microsatellite DNA. Minisatellite DNA consists of repeats with lengths ranging from 9 to 65 base pairs (bp), while microsatellite DNA features repeats of 2 to 8 bp. Microsatellites, more abundant in the genome compared to minisatellites, are widely applied in population genetic analyses. Minisatellites come in two types: multilocus and single-locus. Multilocus minisatellites exhibit tandem repeats ranging from 9 to 65 bp with lengths varying from 0.1 to 7 kilobases (kb). While minisatellite loci are useful for parentage analyses, they are less practical for population genetic studies due to complex mutation patterns. Single-locus minisatellite probes, however, are successful in detecting population genetic variations and find applications in forensics, parentage analysis, understanding mating success, and confirming gynogenesis. Microsatellites, simple repeated sequences highly variable in the genome, serve as valuable markers found approximately every 10 kilobases. These loci are instrumental in genome mapping and population genetic investigations, evolving rapidly at a rate of 10^-3 to 10^-4 mutations per generation. Inherited in a Mendelian fashion, microsatellites are codominant markers, popular for their high polymorphism. Automated genotyping using labeled primers enhances the accuracy and speed of size polymorphism analysis in microsatellites. However, challenges such as null alleles and stutter bands limit their use. While null alleles decrease accuracy in parentage analysis, stutter bands result from slipped strand mispairing during PCR. Microsatellite markers find diverse applications in fisheries and aquaculture, aiding in population genetic structure determination, conservation of biodiversity, phylogenetic and phylogeographic studies, understanding the impact of stocking and hybridization, and forensic identification of individuals.

Single Nucleotide Polymorphisms (SNPs) arise in the genome due to single-point mutations, encompassing insertions/deletions, transitions, or transversions. These mutations result in divergent alleles with alternative bases at specific nucleotide positions, providing valuable insights into intra- and inter-specific diversity patterns. SNPs, regarded as the most prevalent polymorphism in the genome, can be identified through PCR, microarray chips, or fluorescence technology. In fisheries, SNPs are emerging as next-generation markers with widespread applications in population genetics and genomics research.

DNA microarrays, whether small glass microscope slides, nylon membranes, or silicon chips, serve as platforms capable of holding numerous immobilized DNA fragments in a standardized pattern. Reporter probes with known sequences can be matched with DNA from target samples of unknown origin. Microarrays are versatile tools, facilitating

the construction of species-specific DNA probes for subsequent identification purposes. Target sample DNA, labeled with fluorescent molecules, is then hybridized to the microarray DNA. Positive hybridization emits a fluorescent signal, detectable using appropriate fluorescence scanning or imaging equipment.

Expressed Sequence Tags (ESTs), generated through random cloning of cDNA, prove valuable in gene identification and expression analysis. ESTs enable rapid and reliable analysis of genes expressed in specific tissue types under distinct physiological or developmental stages. cDNA microarrays, utilizing ESTs, are effective in identifying differentially expressed genes. Furthermore, ESTs find utility in linkage mapping, enhancing their versatility in various molecular biology applications

Molecular markers in fishery biology investigations

- **To understand inter and intra specific variations**

  The baseline for distinguishing species at the DNA or gene level is determined by the degree of divergence, and decisions should take into account the variability in evolution among taxa. Molecular markers play a crucial role in identifying mixed catches, larval forms, illegally caught endangered and threatened animals, stranded cetaceans, and processed fish products up to the species level. In situations where morphological identification is not feasible, molecular markers provide a reliable alternative. Additionally, these markers can be effectively employed for fish stock characterization and the identification of sub-species.

- **Phylogenetic and Phylogeograhical studies**

  Phylogenetic studies center on historical processes that impact the relationships between species, while phylogeographic studies delve into processes influencing geographical distribution. Mitochondrial DNA markers prove highly effective in both phylogenetic and phylogeographic investigations. Utilizing mitochondrial DNA data enables the reconstruction of the evolutionary history of fish groups, providing crucial insights into historical demography. Moreover, this information aids in deducing conservation units and ecological patterns. Mitochondrial DNA serves as a potent tool for inferring intraspecific phylogenetic patterns in numerous marine fish species.

- **Identification of genetic structure between and within populations**

The crucial aspect of fisheries management and conservation involves identifying the stock structure within fish populations. Stocks represent subpopulations within species, potentially marked by reproductive isolation and distinct physiological or behavioral patterns. Mitochondrial DNA and microsatellite markers are extensively employed to deduce the genetic stock structure of marine fish populations. To attain a comprehensive understanding of subpopulation structure, it is essential to integrate genetic information with morphological and meristic data. Molecular tools also facilitate the identification of subunits within mixed fisheries and determine the origin of stock components, enhancing our ability to manage and conserve fish populations effectively.

- **Genetic tagging/marking**

Marking individual fishes is a valuable practice for monitoring their movement, migration, and assessing population size, especially in mixed fisheries where distinct stocks contribute. Traditional physical tags, being non-heritable, have limitations for generational tracking. However, employing genetic tagging, such as monitoring a rare allele in individuals across generations, proves advantageous. This approach aids in comprehending the impact of hatchery programs on harvest and distinguishing migrants from various regions over an extended timeframe.

- **Forensic investigation**

Molecular markers prove highly efficient in the identification of deceased or stranded fish, as well as preserved or canned items, where morphological identification is impractical. Molecular forensic technologies enable the certification of fishery products and the detection of illicit trading in fish and fishery products. Additionally, molecular tools can be employed to monitor intentional or unintentional releases of fish or organisms into natural waters, enhancing our ability to track and manage such activities.

- **Studying the trophic relationships**

Understanding trophic relationships in an ecosystem is crucial for ecological studies, and acquiring data on diet composition is integral to this endeavor. Identifying diet components up to the species level based solely on morphological features poses challenges, as partial or complete digestion can erase key

morphological characteristics. The extraction of DNA from partially digested samples offers a solution, enabling the study of diet components through the use of molecular markers.

- **Ancient DNA to deduce historical evolutionary relationships**

DNA sequence information can be extracted from preserved samples found in museums, fossil remains, archaeological discoveries, and other unconventional sources, employing various methods. Numerous studies using such approaches have enhanced our comprehension of the evolutionary relationships among different taxa.

- **Applications in aquaculture**

Molecular markers prove highly effective in various aspects of aquaculture, including selective breeding and genetic improvement, identifying quantitative trait loci, detecting inbreeding events, genetically characterizing hatchery stocks, assigning progeny, implementing marker-assisted selection, and comprehending the impacts of ploidy induction and gynogenesis. Moreover, molecular markers find extensive applications in disease diagnosis within aquaculture. PCR-based kits are readily available for the detection of numerous diseases such as white spot syndrome virus (WSSV), infectious pancreatic necrosis virus (IPNV), viral nervous necrosis virus (VNNV), channel catfish virus (CCV), infectious hematopoietic necrosis virus (IHNV), and viral hemorrhagic septicemia virus (VHSV).

References

Allendorf, F.W., Hohenlohe, P.A. and Luikart, G. (2010). Genomics and the future of conservation genetics. Nature Reviews Genetics 11: 697-709.

Askari, G.H., Shabani, H. and Miandare, H.K. Application of molecular markers in fisheries and aquaculture (2019). Scientific Journal of Animal Science, 2(4): 82-88.

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

Cossins, A.R. and Crawford, D.L. (2005). Fish as models for environmental genomics. Nature Reviews Genetics 6:324-340.

Crollius, H.R. and Weissenbach, J. (2005). Fish genomics and biology. Genome Research 15: 1675-1682.

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.

Ding, F., Chu, W., Cui, P., Tao, M., Zhou, R., Zhao, F., Hu, S. and Zhang, J. (2011). ESTBased identification of genes expressed in skeletal muscle of the Mandarin Fish (*Siniperca chuatsi*). Genomics, Proteomics and Bioinformatics 9(1-2): 30-36.

Guo, B., DeFaveri, J., Sotelo, G., Nair, A. and Merila, J. (2015). Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. BMC Biology 13: 19

Hallerman, E.M. (2006). Use of molecular tools for research and improvement of aquaculture stocks. The Israeli Journal of Aquaculture – Bamidgeh, 58(4): 286296.

Hansen, M.M.,Olivieri, I., Waller, D.M., Nielsen, E.E. and The GeM working group. (2012). Monitoring adaptive genetic responses to environmental change. Molecular Ecology 21: 1311-1329.

Helyar, S.J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M.I., Ogden, R., Limborg, M.T., Cariani, A., Maes, G.E., Diopere, E., Carvalho, G.R. and Nielsen, E.E. (2011). Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Molecular Ecology Resources 11(Suppl.1), 123136.

Hemmer-Hansen, J., Therkildsen, N.O. and Pujolar, J.M. (2014). Population genomics of marine fishes: Next – Generation prospects and challenges. Biological Bulletin 227: 117-132.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. and Cresko, W.A. (2010). Population genomics of parallel adaptation in Threespine stickleback using sequenced RAD tags. PLoS Genetics 6(2): 1-23.

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

Nielsen, E.E., Hemmer-Hansen, J., Larsen, P.F. and Bekkevold,D. (2009). Population genomics of marine fishes: identifying adaptive variation in space and time. Molecular Ecology 18: 3128-3150.

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.

Quere, N., Guinand, B., Kuhl, H., Reinhardt, R., Bonhomme, F. and Desmarais, E. (2010). Genomic sequences and genetic differentiation at associated tandem repeat markers in growth hormone, somatolactin and insulin-like growth factor-1 genes of the sea bass, *Dicentrarchus labrax*. Aquatic Living Resources 23: 285296.

Zhu, M. and Zhao, S. (2007). Candidate gene identification approach: progress and challenges. International Journal of Biological Sciences 3(7): 420-427.