Application of molecular markers in fisheries and aquaculture

Gh. Askari, A. Shabani, H. Kolangi Miandare

Department of Fisheries, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran.

Corresponding author; Department of Fisheries, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran.

ARTICLE INFO

Article history:
Received 17 March 2013
Accepted 27 April 2013
Available online 30 April 2013

Keywords:
Genetic diversity
Molecular markers
Microsatellite
Aquaculture

ABSTRACT

Genetic variation in a species enhances the capability of organisms to adapt to changing environment and is necessary for survival of the species. Genetic variation arises between individuals leading to differentiation at the level of population, species and higher order taxonomic groups. The genetic diversity data have varied application in research on evolution, conservation and management of natural resources and genetic improvement programs, etc. Development of Molecular genetic markers has powerful ability to detect genetic studies of individuals, populations or species. These molecular markers combined with new statistical developments have revolutionized the analytical power, necessary to explore the genetic diversity. Molecular markers and their statistical analysis revolutionized the analytical power, necessary to explore the genetic diversity. Various molecular markers, protein or DNA (mt-DNA or nuclear DNA such as microsatellites, SNP or RAPD) are now being used in fisheries and aquaculture. These markers provide various scientific observations which have importance in aquaculture practice recently such as: 1) Species Identification 2) Genetic variation and population structure study of natural populations 3) Comparison between wild and hatchery populations 4) Assessment of demographic bottleneck in natural population 5) Propagation assisted rehabilitation programs. In this review article, we have concentrated on the basics of molecular genetics, overview of commonly used markers and their application along with their limitations (major classes of markers) in fisheries and aquaculture studies.
1. Introduction

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). 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 (Avise, 1994; Linda et al., 1995; Askari and Shabani, 2013). These markers have revolutionized the analytical power, necessary to explore the genetic diversity (Hillis et al., 1996). The conclusion from genetic diversity data has varied application in research on evolution, conservation and management of natural resources and genetic improvement programmes, etc (Ferguson et al., 1995; Liu and Cordes, 2004).

In addition to protein markers, application of DNA markers is finding wide acceptance in population genetics. With DNA markers, it is theoretically possible to observe and exploit genetic variation in the entire genome. Both genomic and mitochondrial DNA is used for varied applications. The commonly used technique are allozyme analysis, types of restriction fragment length polymorphism (RFLP), randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), microsatellite typing, single nucleotide polymorphism (SNP), and expressed sequence tag (EST) markers, etc.

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). 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. The significance of type I markers is becoming extremely important for aquaculture genetics. Type I markers serve as a bridge for comparison and transfer of genomic information from a map-rich species into a relatively map-poor species. 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).

2. Allozyme markers

Analysis of allozyme loci remained one of the most popular approaches in examining population genetics and stock structure questions in fishes (Suneetha, 2000). 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). Isohyets are structurally different molecular forms of an enzyme system with qualitatively the same catalytic function encoded by one or more loci (Markert et al., 1959). Isohyets, which are encoded by different alleles of the same gene locus, are designated as “allozymes” or “alloenzymes” (Starck, 1998). 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). With the strength as codominant marker, ease of use, and low cost, the allozyme markers are popular in population structure and phylogenetic studies, though has limited role in aquaculture genetics.

3. Mitochondrial DNA markers

Mitochondrial DNA (mtDNA) analysis is being increasingly used in recent population and phylogenetic surveys of organisms. Studies of vertebrate species generally have shown that sequence divergence accumulates more rapidly in mitochondrial than in nuclear DNA (Brown, 1985). This has been attributed to a faster mutation rate in mtDNA that may result from a lack of repair mechanisms during replication (Wilson et al., 1985) and smaller effective population size due to the strict maternal inheritance of the haploid mitochondrial genome (Birky et al., 1989). Due to its rapid rate of evolution, mtDNA analysis has proven useful in clarifying relationships among closely related species. Different parts of the mitochondrial genome are known to evolve at different rates (Meyer, 1993). Almost the entire mtDNA molecule is transcribed except for the approximately 1-kb control region (D-loop), where replication and transcription of the molecule is initiated. In general, non-coding segments like the D-loop exhibit elevated levels of variation relative to coding sequences such as the cytochrome b gene (Brown et al., 1993), presumably due to reduced functional constraints and relaxed selection pressure. The 16s rRNA gene in the mitochondrial genome is one of the slowest evolving genes (Meyer, 1993) whereas rapidly evolving regions are control regions (Chow et al., 1997; Gold et al., 1997). Due to non-Mendelian mode of inheritance, the mtDNA molecule is considered as a single locus (Avise, 1994). In addition, because mtDNA is maternally inherited, the phylogenies and population structures derived from mtDNA data may not reflect complete picture of the nuclear genome if gender-biased migration or selection (Birky et al., 1989) or introgression (Chow and Kishino, 1995) exists. Analyses of mtDNA markers have been used extensively to investigate stock structure in a variety of vertebrates including fishes (Avise et al., 1986; Heist and Gold, 1999), birds, mammals and reptiles (Baker and Marshall, 1997; Zink et al., 2000; Menotti-Raymond and O’Brien, 1993; Avise et al., 1998; Shanker et al., 2004).

4. 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). The amplification of genomic DNA by PCR with arbitrary nucleotide sequence primers, RAPD can detect high levels of DNA polymorphisms (Williams et al., 1990; Welsh and McClelland, 1990). 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). 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 (Wirgin and Waldman, 1994).

5. 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 (Morin et al., 1994; Liu and Cordes, 2004). 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 bi-allelic. SNP markers are inherited as co-dominant markers. Several approaches have been used for SNP discovery including SSCP analysis (Hecker et al., 1999), heteroduplex analysis, and direct DNA sequencing. DNA sequencing has been the most accurate and most used approach for SNP discovery. SNPs are not without their limitations, however, might provide marginal additional, or even less, utility in some applications (e.g. relatedness) (Morin et al., 1994).

6. 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; Tautz, 1989; Litt, and Luty, 1989). Abundant in all species studied to date, microsatellite motifs have been estimated to occur as often as once every 10 kb in fishes (Wright, 1993). Microsatellites tend to be evenly 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), 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). 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.

7. New developing markers in fisheries and aquaculture

Various type of DNA markers have been developed, including Allozymes, microsatellites, RAPDs, mt-DNA and SNPs. These markers in fish populations have revealed high levels of genetic variation distributed throughout the fish genome. A recent initiative has been made to accelerate efforts of DNA marker development, genome mapping and species identification. Major progress has been made toward Expressed Sequence Tags (EST) and DNA barcode development in several aquaculture species.

8. Expressed sequence tags (ESTs)

Expressed sequence tags (ESTs) are single-pass sequences generated from random sequencing of cDNA clones (Adams et al., 1991). The EST is used 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 differentially expressed genes to be determined in a systematic way (Wang et al., 1999), in addition to their great value in genome mapping (Boguski and Schuler, 1995). For genome mapping, ESTs are most useful for linkage mapping and physical mapping in animal genomics such as those of cattle and swine, where radiation hybrid panels are available for mapping non-polymorphic DNA markers (Cox et al., 1990). A radiation panel is composed of lines of hybrid cells, with each hybrid cell containing small fragments of irradiated chromosomes of the species of interest. Typically, the cells from species of interest are radiated to break chromosomes into small fragments. The radiated cells are unable to survive by themselves. However, the radiated cells can be fused with recipient cells to form hybrid cells retaining a short segment of the radiated chromosome. Characterization of the chromosomal break points within many hybrid cell lines would allow linkage and physical mapping of markers and genes. In spite of its popularity in mammalian genome mapping (Korwin-Kossakowska et al., 2002; McCoard et al., 2002), 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 et al., 1999). The value of EST resources and applications of bioinformatics in aquaculture genetics/genomics is inevitable, and it is expected that various EST databases will serve as rich sources of genomic information not only for aquaculture geneticists, but also for aquaculture physiologists, immunologists and biotechnologists.

9. DNA barcoding

The principle of conservation biology is the preservation and management of biodiversity. The two major problems to such an endeavor are the difficulty of developing an assessment of this diversity for prioritization of hotspots of species richness (Dobson et al., 1997) and the identification of lineages particularly worthy, or in need, of preservation (Daugherty et al., 1990; Moritz, 2002). Understudied taxa are greatly susceptible to extinction (McKinney, 1995), suggesting there is a conservation penalty for our ignorance. Even there are millions of unidentified and unknown species (Novotny et al., 2002). DNA barcodes, segments of approximately 600 base pairs of the mitochondrial gene cytochrome oxidase I (COI), have been proposed as a fast, efficient, and inexpensive technique to catalogue all biodiversity (Hebert et al., 2003; Hebert et al., 2004). Barcoding is the use of universal polymerase chain reaction (PCR) primers to amplify and sequence an approximately 600-base pair fragment of the COI gene. That portion of sequence is then compared using distance-based algorithms with an existing database of “known” sequences from specimens previously identified by taxonomists. DNA barcodes from a small portion of the mitochondrial genome might seem like an effective and rapid way to assess at least some, perhaps minimal, level of biodiversity. And for groups that are already relatively well known, especially birds and mammals, molecular studies based on barcode sized sequences have revealed cryptic DNA lineages and may be helpful (Hebert et al., 2004).

References

Suneetha, B.K., 2000. 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.


