International Journal of Fauna and Biological Studies
Available online at www.faunajournal.com

IJFBS International Journal of Fauna And Biological Studies

ISSN 2347-2677
IJFBS 2018; 5(4): 83-86
Received: 14-05-2018
Accepted: 15-06-2018

Chandan Haldar
Fish Genetics and Biotechnology Division, ICAR-Central Institute of Fisheries Education, Versova, Mumbai, Maharashtra, India

Raju Ram
Ph.D. Research Scholar, Department of Fish Genetics & Biotechnology, ICAR-Central Institute of Fisheries Education, Versova, Mumbai, Maharashtra, India

Correspondence
Raju Ram
Ph.D. Research Scholar, Department of Fish Genetics & Biotechnology, ICAR-Central Institute of Fisheries Education, Versova, Mumbai, Maharashtra, India

Single nucleotide polymorphism marker and its application in aquaculture: New opportunities and challenges

Chandan Haldar and Raju Ram

Abstract
Genetic variation in a species improves the capability of organisms to adapt to changing environment and is essential for the existence 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 a powerful ability to detect genetic studies of individuals, populations or species. These molecular markers combined with new statistical developments have revolutionized the analytical power, essential to discover the genetic diversity. Molecular markers and their statistical analysis revolutionized the analytical power, necessary to explore the genetic diversity. Different molecular markers, protein or DNA (mt-DNA or nuclear DNA such as microsatellites, SNP or RAPD) are now being used in fisheries. These markers provide various scientific observations which have importance in aquaculture practice recently such as 1) Genetic variation and population structure study of natural populations 2) Species Identification 3) Assessment of demographic bottleneck in natural population 4) Comparison between wild and hatchery populations 5) Propagation assisted rehabilitation programs. In this review article, we have concentrated on the basics of molecular genetics, an overview of commonly used markers and their application along with their limitations (major classes of markers) in fisheries and aquaculture studies. In this review, the current status of SNP genotyping is discussed regarding the mechanisms of allelic discrimination, the reaction formats, and the detection modalities.

Keywords: Single nucleotide polymorphism (SNP), marker assisted selection, next generation sequencing (NGS), high-resolution melting (HRM)

Introduction
A single nucleotide polymorphism (SNP) is a source of variance in the genome. As suggested by the acronym, an SNP ("snip") is a single base mutation in DNA with a usual alternative of two possible nucleotides at a given position. SNPs are the simplest form and the most common source of genetic polymorphisms (Vignal *et al*., 2002) [1] (Fig 1.). There are four major reasons for an increasing interest to use SNPs as DNA markers for genetic analyses. Firstly, they are prevalent and abundant and thus provide more potential markers near or in any locus of attention than other types of polymorphism such as microsatellites. Genome-wide linkage scans tend to employ high density maps of SNPs because both the theoretical and simulation studies, as well as real data application (Liu *et al*., 2004) [2], indicate that SNPs can achieve a superior power to detect and localize linkages in comparison to other makers. All these characteristics of SNPs made it suitable DNA variant for association analyses of quantitative traits in beef cattle. Different types of techniques are available to discover the genetic variation within and among stocks such as RFLP, RAPD, AFLP, microsatellites, etc. However, these markers are difficult to use in high throughput genotyping protocols, because they are present in low density and not evenly distributed in the genome. Alternatively, SNPs markers can be used because of their abundance, even distribution across the genome, high genotyping efficiency, data quality and analytical simplicity (Seeb *et al*., 2011) [3]. The non-synonymous SNPs present in the coding regions of genes are of particular interest because they alter the amino acid sequences of proteins. Such alterations may modify protein function and influence performance traits. A number of SNPs in anonymous genes were reported to be significantly associated with weight gain, growth rate, survival and pathogen resistance in *Litopenaeus vannamei* (Ciobanu *et al*., 2010) [4].

Molecular Markers for Genetic Evaluation

The studied development of molecular genetics laid the basis for genomics that has introduced a new generation of molecular markers for practice in the genetic development of farm animals. These markers provide more accurate genetic information and a better understanding of the animal genetic resources (Samarai & Kazaz, 2015) [5]. Molecular markers have been used in many unlike aspects of genetic managing in aquaculture. Their role in aquaculture has been reviewed by Liu & Cordes (2004) [2]. Molecular markers have been broadly used for identification of strains and species, detection of inter and intra-specific hybridization, parentage and kinship analysis, valuation of parental influence in mass spawning, estimation of effective population size and level of inbreeding, avoiding inbreeding depression, mapping of selective breeding and quantitative trait loci (QTLs). With respect to rehabilitation aquaculture, genetic markers have found a role in the assessment of farmed strains and wild populations, high-quality of donor population, the discovery of genetic changes in hatchery-reared fishes over generations and observing the impact of reared animals after issuing to the wild (Cross, 2000) [6].

Single Nucleotide Polymorphism (SNP)

In 1996, Lander [7] proposed a new molecular marker technology named single nucleotide polymorphism (SNP). It denotes a sequence polymorphism caused by a single nucleotide mutation at an exact locus in the DNA sequence (Yang *et al*., 2013) [8]. Liu & Cordes (2004) [2] defined SNPs as those affected by point mutations giving increase to different alleles containing another base at a given nucleotide position within a locus. This sort of polymorphism includes single base transitions, transversions, insertions and deletions. To be considered an SNP the least frequent allele should have a frequency of 1% or greater (Lander, 1996) [7]. Transitions are the most common (~ 2/3) among all the SNP mutation types. SNPs are becoming the focal point in molecular marker development since they represent the most abundant polymorphism in any organism's genome (coding and non-coding regions). They have a low mutation rate, adaptability to automation and reveal hidden polymorphisms not detected with other markers and methods (Liu & Cordes, 2004) [2]. Theoretically, one 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.

SNP discovery

Several methods have been used for SNP discovery including SSCP analysis, hetero-duplex analysis, real time PCR, microarray and DNA sequencing. DNA sequencing has been the most precise and most used approach for SNP detection and has also developed cost effective with the advent of next generation sequencing (NGS) methods (Vignal *et al*., 2002) [1]. For most accurate detection of SNPs, the Sanger sequencing method is the gold-standard sequencing technology, making it ideal for confirmation of novel variants (Fig. 2). The high throughput Sanger DNA sequencing service gives a high quality resolution of polymorphisms, insertions and deletions in exonic and promoter regions and is useful in resequencing the gene of interest. Resequencing of coding sequences of genes in large populations has previously been shown to be useful for identifying multiple rare variants affecting quantitative traits.

Alignment of multiple sequence fragments originating from different genotypes at the same locus on the genome shows mismatches that are likely to be SNPs or insertions/deletions. Once putative SNPs have been developed for a particular species, SNP genotyping can be accomplished using several methods, including single strand conformation polymorphism (SSCP), allele-specific amplification (ASA), denaturing gradient gel electrophoresis (DGGE), Taq-Man probes, high-resolution melting (HRM) analysis, denaturing high-performance liquid chromatography (DHPLC), SNP chips and DNA sequencing (Kinoshita- Kikuta *et al*., 2007) [9]. However, these genotyping methods involve complicated processes, specific apparatus, expensive reagents, and/ or a high skill level (Kinoshita-Kikuta *et al*., 2007) [9]. If a sufficiently large number of validated SNPs are available, SNP chips can be prepared for simple, rapid, high-throughput, accurate and low-cost SNP genotyping.

SNP Studies in Aquatic Animals

Fridjonsson *et al*. (2011) [10] detected mtDNA SNPs in Atlantic salmon using the high throughput sequencing technology (GSFLX 454) by pooling 546 individuals in 16 groups. They identified a total of 216 SNPs and 2 indels that were validated and mapped onto the *S. salar* mitochondrial genome. Of these, 108 markers (107 SNPs and one indel) were novel. Liu *et al*. (2011) [11] undertook transcriptome sequencing and analysis for both channel catfish and blue catfish using Illumina NGS platform to develop large numbers of high-quality SNPs. Such a method is most effective when coupled with existing EST resources generated using outmoded sequencing approaches because the reference ESTs facilitate the effective meeting of the expressed short reads. When many individuals with different genetic backgrounds are used, this method is very effective for the identification of SNPs. Mohindra *et al*. (2012) [12] mined microsatellite repeats and SNPs from the spleen ESTs of *C. batrachus*. From a total of 1,698 unique ESTs, 54% could be functionally annotated by similarity searches and 221 SSRs were identified. A total of 23 contigs covering 3 or more ESTs were found to contain 31 SNP loci. Nine ESTs with SSRs and/or SNPs identified in this study have been earlier stated to be related to diseases in human and animals. Table 1. showing the different studies conducted on SNPs in fishes.

Jung *et al*. (2011) [13] identified the candidate genes and SNPs from 454GS-FLX transcriptome data in muscle ovary and testis tissue. Of the 834 SNPs detected, 555 were putative transitions (Ts) and 279 were putative transversions (Tv), giving a mean Ts: Tv ratio of 1.99 : 1.00 across the transcriptome. The SNP types A↔ G and C ↔ T were most common and SNP densities varied among genes. Blanck *et al*. (2013) [14] reported the discovery of 13 synonymous SNPs within a fragment of Hsc70 gene in *M. amazonicum*. The *M. rosenbergii* Hsc70 gene sequence was used as a reference in this study.

Advantages and limitations of SNP markers

The most recently designated type of marker is the single nucleotide polymorphism (SNP, pronounced "snip"). As the name suggests, these are variations at a single base on the chromosome. For example, on some chromosomes a locus might have a C, while on other chromosomes the same locus might have a T. These are the most common markers, with at

least three million already labeled, and seem to occur diagonally the entire genome. As with RFLPs, there are nearly always only two alleles at an SNP locus. Separately they suffer the same problem as RFLPS of not being valuable in many of the families. They are being used broadly now because they are very easy to genotype, are very common (happening at least ten times more frequently than the other types of markers) and thus can be used in mixture with each other.

Advantages

- PCR products can be very small:–Markers will work with extremely degraded DNA samples.
- SNPs are more common in the genome. SNPs are found in coding and (mostly) non coding regions and they occur with a very high occurrence about 1 in 1000 bases to 1 in 100 to 300 bases. The richness of SNPs and the ease with which they can be measured make these genetic variations significant. SNPs close to particular gene acts as a marker for that gene. SNPs in coding regions may alter the protein structure made by that coding region.
- It may possibly multiplex hundreds or thousands on one chip.
- Sample processing may be completely automated.
- No stutter products.

Disadvantages

- Its PIC (polymorphic information content) is lower than microsatellite markers because, it is a bi-allelic. Each marker is less informative.
- Therefore have to genotype many more SNPs to get the same level of information about DNA sample.
- Mixture interpretation is more difficult
- Multiplexes don't actually work yet. Currently uses more of the DNA sample than STRs use.

Application in fisheries

SNP markers are perhaps most influential for genome mapping and identification of candidate genes for QTL, but their discovery needs great economic investment. Effectual genotyping also requires exclusive equipment. Mass spectrometers cost over US$300,000, pyrosequencers cost about US$100,000, and quantitative PCR equipment costs about US$50,000. Additionally, processing currently costs about US$1 per genotype for SNPs. In spite of the cost, it is foretold that SNPs will be the future marker of choice in biotechnology-related industries due to their nearly limitless power and adaptability to automation. While it is not so sure that SNPs will become popular in aquaculture genetics due to the financial limits listed above, it is almost certain that RFLPs and allozymes represent markers of the past (Liu, 2004) [2]. Since most, if not all, production and performance traits are controlled by multiple genes and therefore inherited as quantitative traits, analysis of their associated quantitative trait loci (QTL) is a very important part of aquaculture genetics/ genomics (Liu *et al*, 2004) [2]. QTLs are largely unidentified genes that affect performance traits (such as growth rate and disease resistance) that are vital to breeders. Relative chromosomal positions of QTL in a genome can be identified in a two-step process that begins by making a genetic linkage map. Genetic linkage maps are made by allocating (mapping out) polymorphic DNA markers (such as microsatellites, SNP or AFLPs) to chromosome configurations based on their separation relationships. The SNPs are linked to QTLs form the basis of marker assistance selection (MAS) in a selective breeding program for successful aquaculture production.

Future applications of DNA markers in aquaculture genetics

In addition to genome mapping and the other applications debated in this review, DNA markers are likely to show useful in many other features of aquaculture. The growth and application of DNA marker technologies already ongoing in other areas such as molecular systematics, population genetics, evolutionary biology, molecular ecology, conservation genetics, and seafood security monitoring will undoubtedly impact the aquaculture industry in unforeseen ways. Already, lessons cultured from studies in population and conservation genetics are changing the very role that hatcheries and aquaculture play for increase and restoration of wild fish stocks such as salmon and trout. Advances in aquaculture genomics are also likely to disturb other areas using molecular markers as well. Though it may take some time to implement marker-assisted selection in aquaculture, the techniques of genome mapping and QTL analysis used to support MAS will finally also be used to classify and clone genes that could show to be economically significant outside of the aquaculture arena, and find applications in medicine and other bio-related industries.

Table 1: Different studies on SNPs in fish

Author (Year)	Species	Method used
Andreassen *et al*. (2010) [15]	Atlantic salmon (Salmo salar)	Full length transcripts were sequenced by Sanger's method
Souche *et al*. (2007) [16]	European sea bass	Expressed sequence tag sequencing
Fridjonsson *et al*. (2011) [10]	mtDNA SNPs in Atlantic salmon	Using the high throughput sequencing technology (GSFLX 454) by pooling 546 individuals in 16 groups
Liu *et al*. (2011) [11]	Channel catfish and blue catfish	Using the Illumina NGS platform
Jung *et al*. (2011) [13]	*Macrobrachium rosenbergii*	454GS-FLX transcriptome sequencing

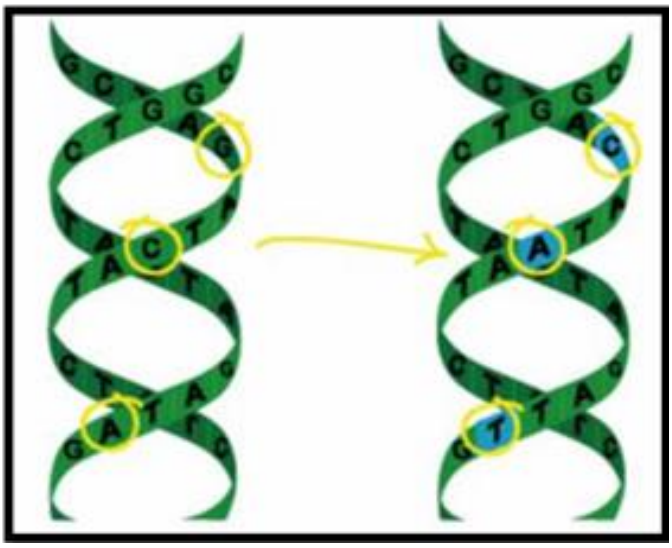

Fig 1: 3 SNPs are found between the two DNA fragments which are G/C, C/A, and A/T respectively.

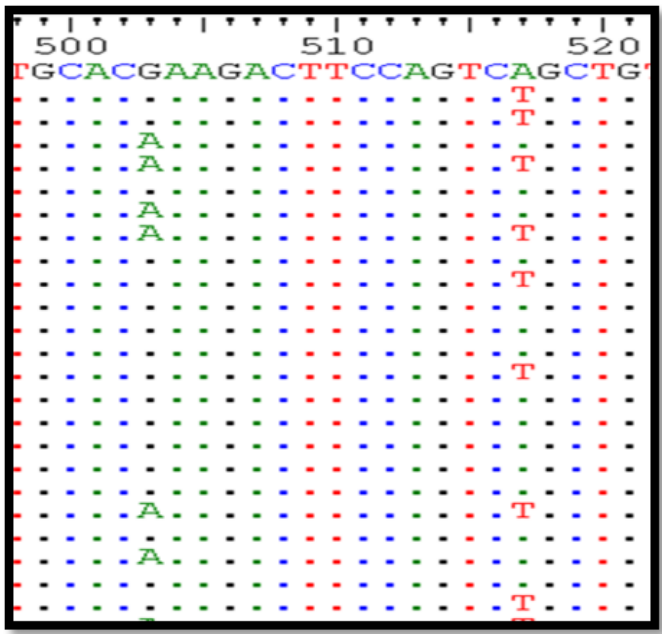

Fig 2: Detection of SNPs in 503(G/A) and 517(A/T) using g DNA sequences in Bio Edit multiple alignment tool (Clustal W).

References

1. Vignal A, Milan D, Cristobal M, Eggen A. A review on SNP and other molecular markers and their use in animal genetics. Genetics Selection Evolution. 2002; 34:275-305.
2. Liu ZJ, Cordes JF. DNA marker technologies and their applications in Aquaculture Genetics. Aquaculture. 2004; 238:1-37.
3. Seeb. Seeb J, Carvalho G, Hauser L, Naish K, Roberts S, Seeb L. Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Molecular Ecology Resources. 2011; 11:1-8.
4. Ciobanu DC, Bastiaansen JWM, Magrin J, Rocha JL, Jiang DH, Yu N *et al*. A major SNP resource for dissection of phenotypic and genetic variation in Pacific white shrimp (*Litopenaeus vannamei*). Animal genetics. 2010; 41:39-47.
5. Samarai FR, Al-Kazaz AA. Applications of Molecular Markers in Animal Breeding: A. American Journal of Applied Scientific Research. 2015; 1:1-5.
6. Cross TF. Genetic implications of transaction and stocking of fish species with particular reference to Western Australia. Aquaculture Res. 2000; 31:83-94.
7. Lander ES. The new genomics: global views of biology. Science. 1996; 274:536.
8. Yang W, Kang X, Yang Q, Lin Y, Fang M. Review on the development of genotyping methods for assessing farm animal diversity. Journal of animal science and biotechnology. 2013; 4:1.
9. Kinoshita E, Kinoshita-Kikuta E, Koike T. A single nucleotide polymorphism genotyping method using phosphate-affinity polyacrylamide gel electrophoresis. Analytical biochemistry. 2007; 361(2):294-298.
10. Fridjonsson O, Olafsson K, Tompsett S, Bjornsdottir S, Consuegra S, Knox D *et al*. Detection and mapping of mtDNA SNPs in Atlantic salmon using high throughput DNA sequencing. BMC genomics. 2011; 12:179.
11. Liu S, Zhou Z, Lu J, Sun F, Wang S, Liu H *et al*. Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array. BMC genomics. 2011; 12(1):53.
12. Mohindra V, Singh A, Barman AS, Tripathi R, Sood N, Lal KK. Development of EST derived SSRs and SNPs as a genomic resource in Indian catfish, *Clarias batrachus*. Mol. Biol. 2012; 39:5921-5931.
13. Jung H, Lyons RE, Dinh H, Hurwood DA, McWilliam S, Mather PB. Transcriptomics of a giant freshwater prawn (*Macrobrachium rosenbergii*): de novo assembly, annotation and marker discovery. PloS one. 2011; 6(12):e27938.
14. Blanck DV, Valenti WC, Freitas PD, Galetti J. Isolation and characterization of SNPs within HSC70 gene in the freshwater prawn *Macrobrachium amazonicum*. Conservation Genetics Resources. 2013; 5:631-633.
15. Andreassen R, Lunner S, Høyheim B. Targeted SNP discovery in Atlantic salmon (Salmo salar) genes using a 3'UTR-primed SNP detection approach. Bmc Genomics. 2010; 11(1):706.
16. Souche EL, Hellemans B, Van Houdt JKJ, Canario A, Klages S, Reinhardt R *et al*. Mining for single nucleotide polymorphisms in expressed sequence tags of European sea bass. Journal of Integrative Bioinformatics. 2007; 4(3):158-167