Organizing committee:

Prof (Dr.) Mohan Joseph Modayil
Director, CMFRI, Kochi

Dr. P.C. Thomas
Winter School Director

Co-ordinators:

Dr. R. Paul Raj, Head, PNPD
Dr. K.C. George, Principal Scientist
Dr. P. Jayasankar, Senior Scientist
Dr. D. Noble, Senior Scientist

WINTER SCHOOL

ON

Recent Advances in Mariculture Genetics and Biotechnology

4th to 24th November 2003

Course Manual

INDIAN COUNCIL OF AGRICULTURAL RESEARCH
CENTRAL MARINE FISHERIES RESEARCH INSTITUTE
P.B. No. 1603, Tatapuram P.O.,
Kochi – 682 014

DNA MARKERS FOR FISHERIES APPLICATIONS

P. Jayasankar

Central Marine Fisheries Research Institute, Cochin

Introduction

Over the past few years, application of molecular genetic markers in fisheries has increased dramatically due to the advances in DNA sequencing, data analysis and PCR and it has been feasible to tackle several issues including population genetics, broodstock development, fish health management, transgenics, genetic diversity, conservation and genomics. Molecular markers are polymorphic DNA or protein sequences that can be used to identify a chromosomal region. The molecular markers blended with the PCR technology have become the central tool in many areas of fisheries research. Excellent reviews on the role of molecular genetic markers in fisheries are available (Carvalho and Pitcher, 1994; O'Reilly and Wright, 1995; Ferguson and Danzmann, 1998). Molecular markers are typically unaffected by environmental conditions unlike meristic and morphometric characters. There are two types of markers based on their origin: *Protein markers* and *DNA markers* and the present paper deals with the latter.

Advantages of DNA markers over protein markers

DNA markers are based on the polymorphisms detected at the DNA level. Polymorph DNA markers serve as landmarks or anchor loci for identification and analysis of new loci/genes in the genome. The methods involving direct examination of DNA are strikingly different from allozyme techniques in the following respects (a) the range of potential genetic markers that can be assayed are almost limitless and (b) unlike those using allozymes, DNA researchers use a plethora of different analytical techniques and methods of detection. Among the several advantages of DNA-level markers over protein markers include requirement of only a small amount of tissue, which could be ethanol-preserved or frozen for DNA extraction (DNA can be extracted even from formalin-preserved tissues) and the availability of innumerable potential markers. For protein markers, more amount of tissue is required, non-invasive sampling is not possible and tissue should be fresh or well frozen.

DNA polymorphisms

DNA markers are based on polymorphism detected at the DNA level. Polymorph DNA markers serve as landmarks or anchor loci for identification and analysis of new loci/genes in the genome. Polymorphism information content (PIC) is the single most important characteristic of a marker and is calculated from the allelic frequencies in the population. A PIC value greater than 0.5 is considered as highly informative, a PIC value between 0.25 and 0.5 indicates a reasonably informative marker while markers with PIC smaller than 0.25 are only slightly informative.

Categories of DNA-level markers

A discussion on the categories of DNA, based on function, structure, location etc. is given elsewhere (Jayasankar, 1997) and will not be included in the present paper. Based on their applications, DNA-level markers can be broadly put into two categories, type I and type II. Type I markers are the coding gene loci conserved across the species and are normally monomorphic or slightly polymorphic, often with two alleles. Restriction Fragment Length Polymorphism (RFLP) is an example of type I marker and contribute considerably for mapping of genes. Type II markers are highly polymorphic (PIC>0.6) and useful for population genetics and molecular taxonomy analyses. Random Amplified Polymorphic DNA (RAPD), Amplified Fragment Length Polymorphisms (AFLP) and Microsatellites are examples of type II markers.

Restriction Fragment Length Polymorphisms (RFLP): Restriction enzymes are endonucleases, which occur in a variety of prokaryotes and their natural function is to destroy foreign DNA molecules by recognizing and cutting specific DNA sequence motifs typically consisting of four to six bases. Each enzyme has a particular recognition sequence, and the host bacteria usually protect their own DNA from being cut, by methylating this sequence.

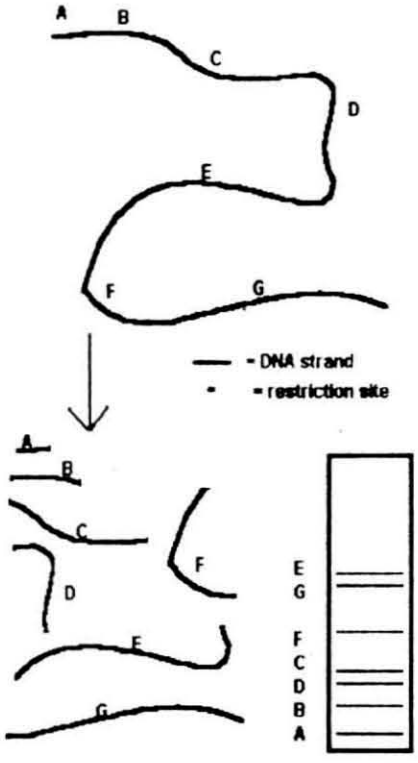

Fig. 1 Diagramatic representation of RFLP

Analysis of RFLPs for evaluating DNA sequence variation is widely used, including fisheries field. Briefly, genomic DNA is extracted, digested with restriction enzymes (a large number of them are now available commercially) and separated by electrophoresis on a gel. The gel is blotted to a nylon membrane and hybridized by a labeled probe, which is a piece of DNA. RFLP probes are locus-specific and easy to screen co-dominant markers, hence widely used for genome mapping. They can be generated from either genomic library or cDNA library.

There are two approaches to study RFLPs in the cytoplasm. The first is to extract mitochondrial DNA (mtDNA) separately from nuclear/genomic (nDNA) and digest them with restriction endonucleases, resolve it in gel and stain. The second strategy is to isolate and digest the total DNA of the organism, followed by electrophoresis and southern blotting. Polymorphisms can be visualized using specific mtDNA probes.

The advantages of RFLPs are:

- Highly polymorphic - many alleles may be present in a population for a single locus (This is an important concept - polymorphism refers to the degree of variation in the

population under consideration. Any individual can have, of course, a maximum of 2 alleles)

- Co-dominant inheritance
- Many loci can be established

The disadvantages of RFLPs are:

- The technique is laborious
- Time-consuming
- Expensive
- Usually uses isotope

Random Amplified Polymorphic DNA (RAPD): Random-amplified polymorphic DNA (RAPDs) involves the use of a single 'arbitrary' primer (purchasable from commercial companies) in a PCR reaction and result in the amplification of several discrete DNA products (Williams *et al.*, 1990; Welsh and McClleland, 1990). Each product is derived from a region of the genome that contains two short segments in inverted orientation, on opposite strands those are complementary to the primer and sufficiently close together (within 2.5-3.0 kb) for the amplification to work. In RAPDs, the amplification products are separated on agarose gels in the presence of ethidium bromide and visualized under ultraviolet light. It is now widely recognized that to obtain reproducible band profiles on the gels it is absolutely essential to maintain consistent reaction conditions.

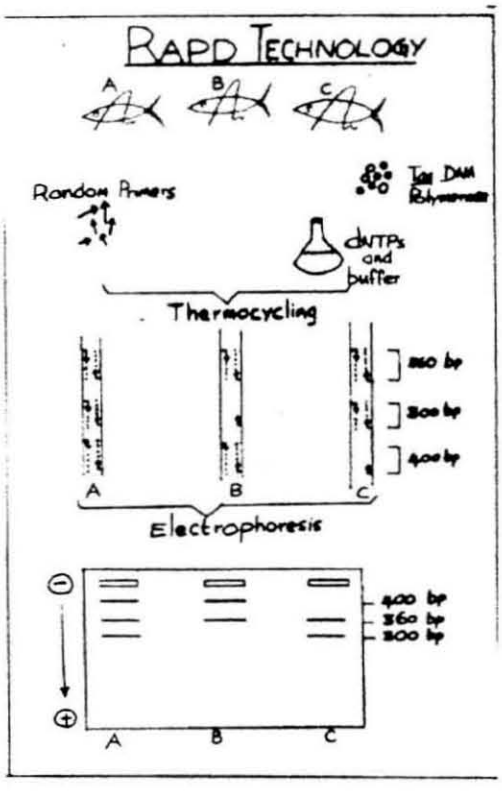

Fig. 2 Basic steps involved in RAPD

Numerous studies have reported the separate effects of altering different parameters, ratio of template DNA primers, concentration of *Taq* DNA polymerase and Mg concentration on the bands obtained (Smith, 2003). A corollary of these experiments is that RAPD profiles should be reproducible among laboratories provided that all details of the reaction conditions are standardized and strictly adhered to.

Advantages of RAPDs are:

- Rapid, simple, relatively inexpensive assay
- Many loci can be identified quickly.
- The assay can be automated.

Disadvantages of RAPDs are:

- Polymorphism is typically dominant in nature.
- Low allelic polymorphism
- Inconsistency of results

Amplified fragment length polymorphism (AFLP): AFLP is another PCR-based method which first involves restriction digestion of the genomic DNA (Vos *et al.*, 1995). Adapters are ligated

to the ends of the restricted fragments and either a pre-selection step performed using magnetic beads followed by a round of selective PCR, or two selective rounds of PCR amplification are applied. The number and composition of the selective nucleotides used as well as the complexity of the genomic DNA determine the number of resulting amplified fragments.

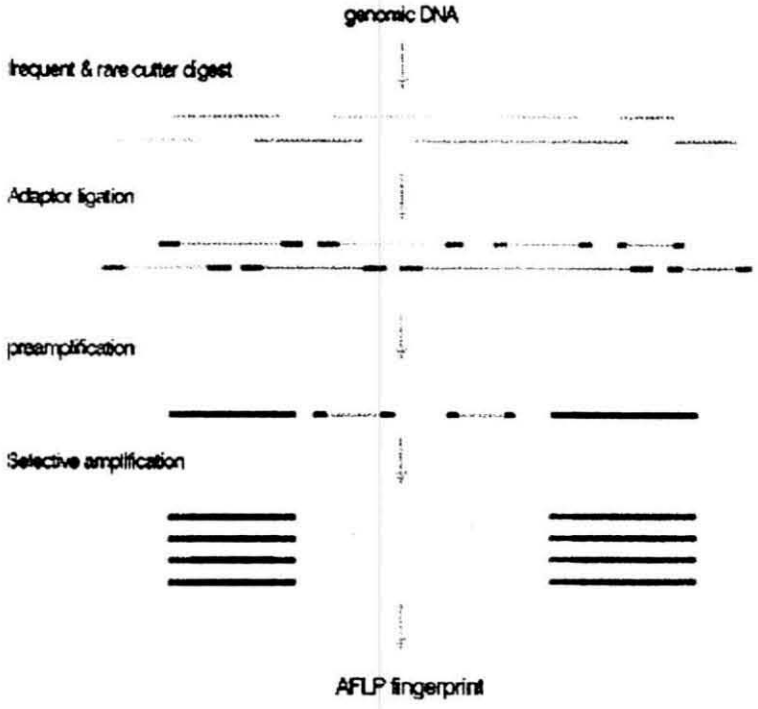

Fig. 3 Processes involved in AFLP fingerprinting

The amplified products are separated on a sequencing gel and can be visualized using radioactive or fluorescent labelling. All the current evidence suggests that AFLPs are as reproducible as restriction fragment length polymorphism (RFLP). They should therefore be highly suited to network experiments.

The advantages of AFLPs are:

- Many more bands, and so potentially many more polymorphisms, are identified than with RFLPs or even RAPDs.
- As with RAPDS, only a small amount of template DNA is needed and no probe hybridization is needed.

- Banding patterns are more consistent than with RAPDs

The disadvantages of AFLPs are:

- The method is labor intensive and requires isotope.
- The technology is proprietary.
- Bands are still scored as present or absent (i.e. dominant or recessive)

Microsatellites: Microsatellites or simple sequence repeats (SSRs) are usually one to four nucleotides long repeat units and are highly mutable loci, which may be present at many sites in a genome, occur as often as once every 10 kbp, and hence have an overall abundance on the order of 10^4-10^5 per genome (Wright and Bentzen, 1994). Some marine fishes and invertebrates exhibit 10-60 alleles per locus with high heterozygosity rates. As the flanking sequences at each of these sites may be unique, once SSR loci are cloned and sequenced, primers can be designed to the flanking sequences. The resultant sequence tagged microsatellite usually identifies a single locus, which because of the high mutation rate of SSRs, is often multi-allelic. Alleles which differ in many base pairs of length can be resolved on agarose gels but often SSRs are visualised on sequencing gels where single repeat differences can be resolved and, thus, all possible alleles detected.

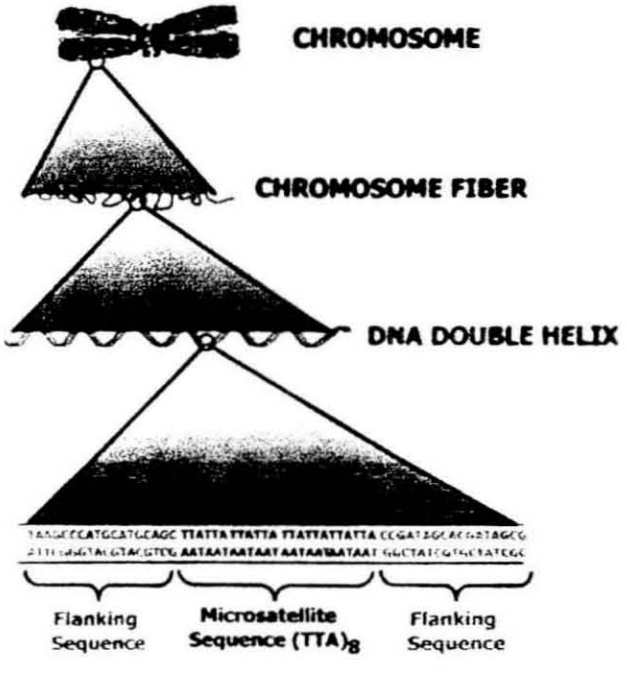

Fig. 4 Diagramatic representation of microsatellite sequences

SSRs provide highly informative markers because they are co-dominant (unlike RAPDs and AFLPs) and generally highly polymorphic. The nature of the PCR-based assay used in their amplification and detection (i.e. the use of specifically designed primers based upon the flanking sequences) suggests that they should be highly reproducible between laboratories.

Advantages of SSRs are:

- Microsatellites are easy to detect via PCR
- They generally display a great deal of polymorphism.
- They are co-dominant in nature

The disadvantages of SSRs are:

- Initial identification requires laborious screening of libraries or some other method of obtaining sequence information so that primers can be designed
- Often stutter bands appear

Mitochondrial genome: Mitochondria are cytoplasmic organelles responsible for respiratory function in eukaryotic cells. The mitochondrial genome is circular double stranded DNA with a size of 16-20 kb and containing about 35 genes. Unlike nuclear genome, mtDNA is haploid, with maternal inheritance. Several regions, particularly control region and cytochrome B gene have high evolution rate and can be used as markers for population and evolutionary genetic studies.

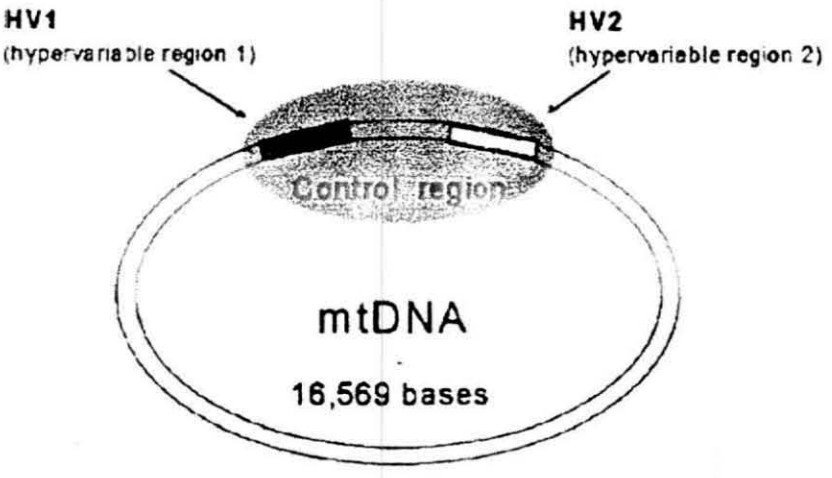

Fig. 5 Hypervariable regions of mitochondrial genome

Remarks

The recent innovations in molecular technology have increased the potential for molecular markers to provide useful information in fisheries management and aquaculture. Markers such as microsatellites have provided increased resolution power to answer stock questions in species having relatively low genetic variation. Other advances include determination of family structure, location of useful genes, etc. The specific requirement and available resources should decide choice of suitable genetic marker.

References

Carvalho, G.R. and T.J. Pitcher 1994. Molecular genetics in Fisheries. *Rev. Fish. Biol. Fish.*, 4: 269-399.

Ferguson, M.M. and R.G. Danzmann 1998. Role of genetic markers in fisheries and aquaculture: useful tools or stamp collecting? *Can. J. Fish. Aquat. Sci.*, 55: 1553-1563.

Jayasankar, P., 1997. DNA fingerprinting fosters genetic research in fisheries and aquaculture. *Fishing Chimes* 17(2): 11-15.

O'Reilly, P and J.M. Wright 1995. The evolving technology of DNA fingerprinting and its application to fisheries and aquaculture. *J. Fish Biol.*, 47 (Supplement A): 29-55.

Smith, P.J., 2003. Genetic analysis: Random Amplified Polymorphic DNA (RAPD). In: *Stock identification methods*, Academic Press. 21 pp.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J. Pelman, M. Kuiper and M Zabeau 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.*, 23:4407-4414.

Welsh, J., and M. McClelland (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res* 18: 7213-7218.

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski, and A. Tingey, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res* 18: 6531-6535.