**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

# GENE CLONING – TECHNIQUES AND STRATEGIES

***P. Jayasankar***

*Central Marine Fisheries Research Institute, Cochin*

## Introduction

Genetic information is coded in DNA found in the nucleus and mitochondria of animal cells and chloroplast of plant cells. The flow of genetic information from DNA to RNA and subsequently to polypeptides or proteins forms the central dogma of molecular biology. From the complex pool of DNA how can one isolate and study a specific sequence, or in other words, a gene of interest? The technology of gene cloning allows the amplification and recovery of a specific fragment from the highly complex DNA pool so that it can be comprehensively studied with regard to its structure, function, regulation, expression, etc. This technology has been immensely useful in the areas of medicine, agriculture and industry. For example, it is now possible to produce proteins to be used as diagnostics and therapeutic agents through recombinant DNA technology. The term must be distinguished from the natural DNA recombinants that result from crossing over between homologous chromosomes in eukaryotes and prokaryotes. It is used in the sense of unnatural union of DNAs from non- homologous sources, usually from different organisms. The alternative term chimeric DNA is derived from the Greek mythological monster Chimera, a combination of parts of different animals.

In practice, DNA cloning can be achieved in two different ways: cell based cloning and PCR based cloning.

## Cell based cloning

### *Principle and approaches*

Cell-based DNA cloning strategy is the most popular method in which the novel combination of DNA from a 'donor' organism and a 'host' organism would take place through a 'vector'. The basic procedure involves the following steps:

- Extraction of DNA from the donor organism ('foreign DNA') and vector (such as plasmids, bacteriophages, etc.)
- Cutting it up into several fragments using restriction endonucleases
- Inserting the cut donor DNA into cut vectors
- Introducing the recombinant DNA into host cell (e.g. *E. coli*) and propagation of host cells in suitable culture medium.

Several clones of host cells are grown in the medium containing several fragments of the donor DNA and it will be the most important job to select the clone containing the gene of interest. Figure 1 gives a brief out line of procedure.

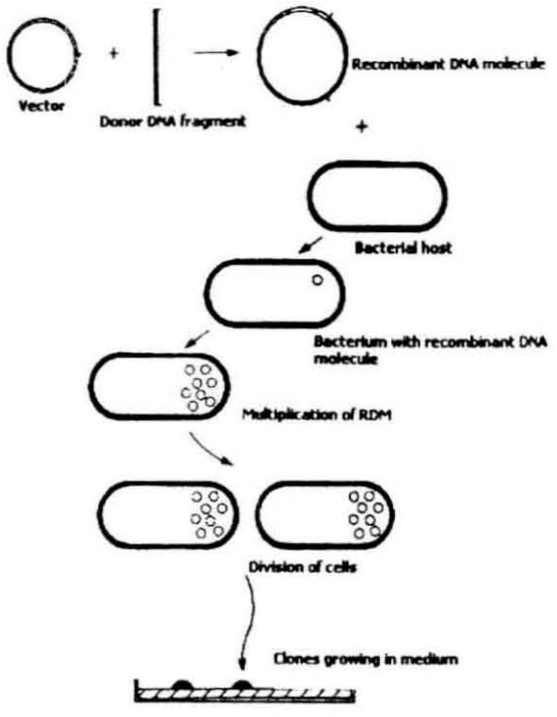


**Fig. 1 The basic steps involved in gene cloning**

*Extraction of DNA from the donor organism ('foreign DNA')*: Genomic DNA from the donor would be extracted using standard protocols, the most common being the proteinase K digestion followed by phenol-chloroform extraction (Sambrook *et al.*, 1988). Vector, such as plasmid carrying genes for resistance to antibiotics can be separated from the bacterial chromosomal DNA. Because differential binding of ethidium bromide by the two DNA species makes the circular plasmid DNA heavier than the chromosomal DNA, the plasmids form a distinct band on centrifugation in a cesium chloride gradient and can be separated.

*Cutting it up into several fragments using restriction endonucleases*: Restriction enzymes act as molecular scissors, cutting DNA at specific sites situated in the recognition sites. Restriction sites are not relevant to the function of the organism, nor would they be cut in vivo because most organisms do not have restriction enzymes. There are three categories of restriction endonucleases, type I, type II and type III out of which only type II find application in molecular biology. The recognition sites generally vary from 4-8 bp and are mostly palindromic in nature. This means, the recognition sequences are same on both the strands when read 5′ 3′ direction. Some enzymes make a 'staggered cut' within the recognition sequence resulting in single stranded sticky ends conducive to the formation of recombinant DNA, while cut of others result in blunt ends.

*Inserting the cut fragments into a cut vector*: Figure 2 shows the method of generating a chimeric DNA; the vector is a plasmid that carries one *Eco* RI restriction site and the donor molecule also has *Eco* RI sites. When the two populations of restriction enzyme-digested

fragments are pooled, they unite due to the presence of sticky ends. At this stage, although duplexes are formed to generate a population of chimeric molecules, the sugar-phosphate backbones are still incomplete at two positions at each junction. These fragments can be linked permanently by the addition of the enzyme DNA ligase, which creates phosphodiester bonds at the joined ends to make a continuous molecule.

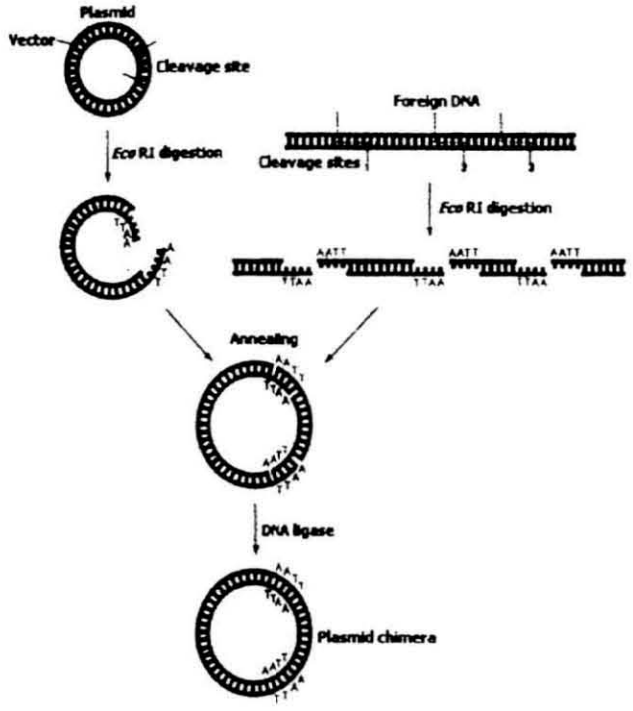


**Fig. 2 Method for generating a chimeric DNA plasmid containing genes derived from foreign DNA**

*Introducing the recombinant DNA into host cell (e.g. E. coli) and propagation of host cells in suitable culture medium*: Recombinant plasmid DNA is introduced into competent host cells by transformation. Once in the host cell, the vector will replicate in the normal way, but now that the donor DNA is a part of its length, the donor DNA is automatically replicated along with the vector. Each recombinant plasmid that enters the cell will form multiple copies of itself in that cell. Subsequently many cycles of division will occur, and the recombinant vectors will undergo more rounds of replication. The resulting colony of bacteria will contain billions of copies of the single donor DNA insert. This set of amplified copies of the single donor DNA fragment is the DNA clone (Fig. 3).

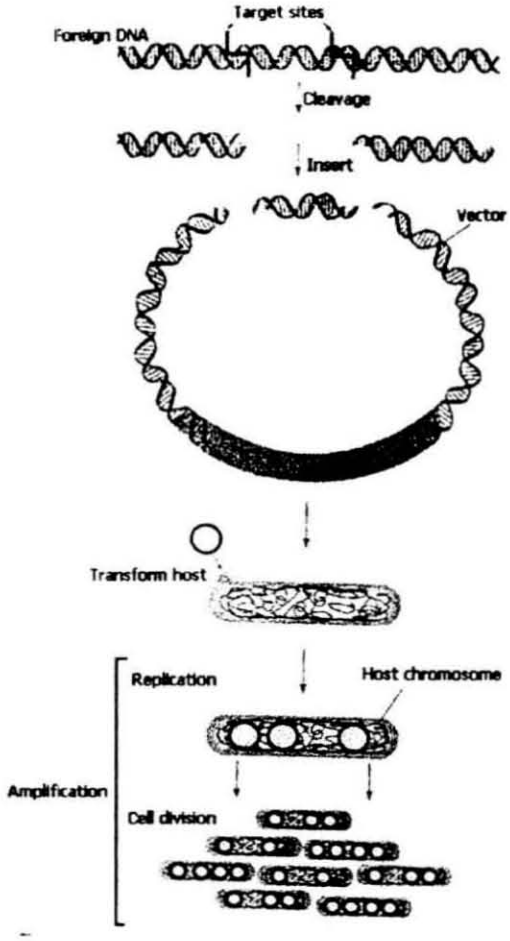


**Fig. 3 Amplification of recombinant DNA in host cell**

### *Choice of vectors*

Some of the basic essential features of vectors used in recombinant DNA technique are as follows

- Must be relatively small molecules for convenience of manipulation.
- Must be capable of prolific replication in a living cell, thereby enabling the amplification of the inserted donor fragment
- Presence of multiple/unique restriction sites for insertion of DNA to be cloned

Important cloning vectors are plasmids, bacteriophage, cosmids and single stranded phages, such as M13. Some of the important expression vectors (those in which cloned gene can be transcribed and translated into proteins) include YACs (Yeast Artificial Chromosome), in which upto 1000 kb fragment can be cloned and BACs (Bacterial Artificial Chromosome), which can clone upto 300 kb fragments. Excellent descriptions of these vectors appear in reviews by Griffiths *et al* (1996) and Brown (1990).

### *DNA library*

The specific approach to clone a gene depends to a large degree on the gene in question and on what is known about it. Generally the procedures start with obtaining a large collection of clones made from the original DNA sample. The collection of clones is known as a DNA library. This step of making a library is sometimes referred, as 'shotgun approach' because the experimenter clones a large sample of fragments hoping that one of the clones will contain a "hit" – the desired gene.

Two basic methods are available for construction of a DNA library, depending the starting DNA: *Genomic DNA library* and *cDNA library*. In case of genomic DNA library, the starting material is the genomic DNA of various sources. Typically, pure genomic DNA is cleaved by restriction endonucleases, preferably 4 bp cutters such as *Mbd* or *Sau* 3A in order to achieve fragmentation of the genome. The optimal fragmentation can be achieved through partial digestion involving low enzyme concentration and shorter incubation time. These fragments after purification and modification can be inserted into suitable vectors for cloning. Though any vector can be used for this purpose, bacteriophage λ vectors are often the choice because inserts up to 25 kb can be easily cloned in this vector (Fig. 4).

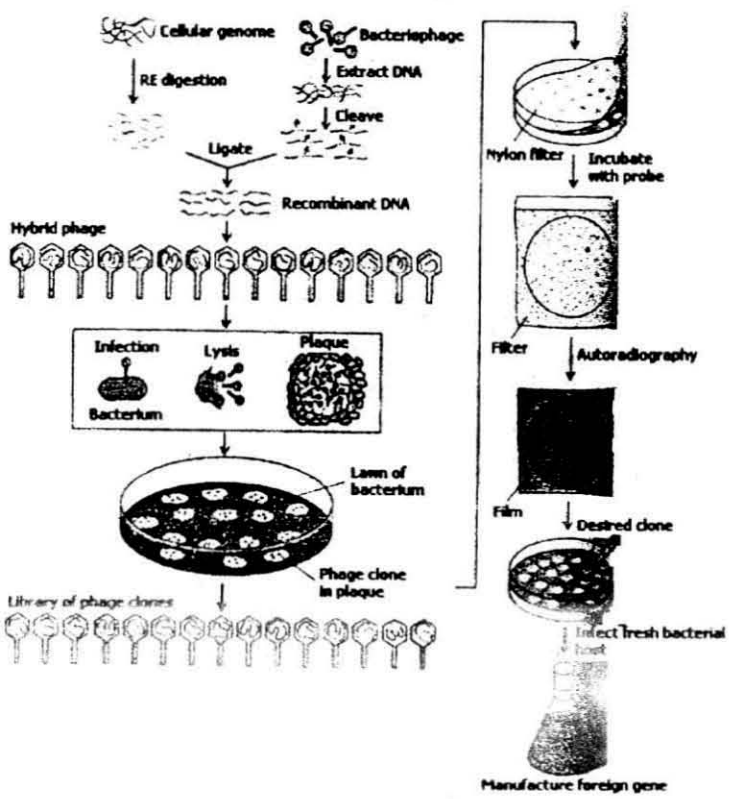


**Fig. 4 Construction of a genomic DNA library from λ phage**

In case of cDNA library, the starting material is usually total RNA from a specific tissue. Using the total RNA as template and enzyme reverse-transcriptase cDNAs are produced (Fig. 5). Messenger RNA too, is often used as template. The resulting double stranded cDNAs are cloned into suitable vectors.

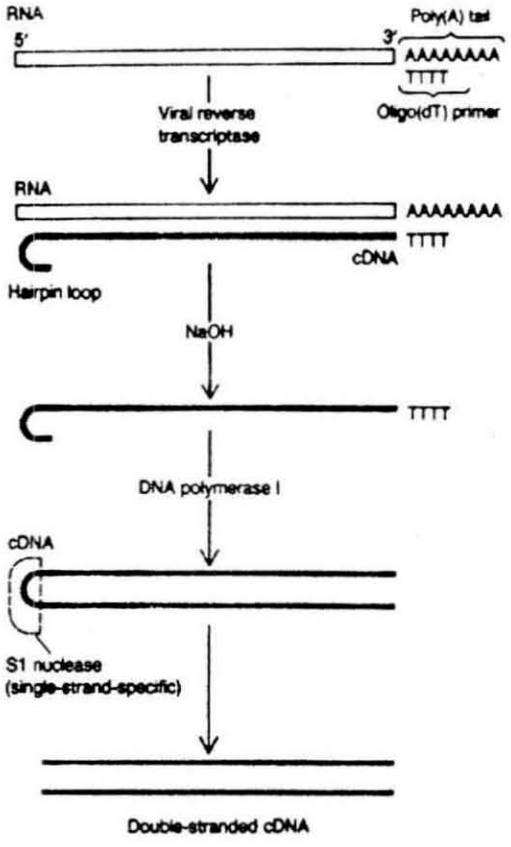


**Fig. 5 Synthesis of double stranded cDNA from mRNA**

The choice between genomic DNA and cDNA depends on the situation. If the gene sought were active in a specific type of tissue in the fish or shellfish, then it would make sense to use that tissue to prepare mRNA to be converted to cDNA, and then make the cDNA library. Because it is made from mRNA, the cDNA would lack in regulatory sequences and introns and thus can be translated into functional proteins in bacteria, a very important feature when cloning and manipulating eukaryotic genes in bacterial hosts. On the other hand, though the genomic library is bigger, they do have genes in native form, including introns and regulatory sequences. But the disadvantage is location of a specific gene from a genomic library is far too cumbersome.

### *How to find a specific clone from a library?*

The library, which might contain as many as hundreds of thousands of cloned fragments, must be screened in order to find the recombinant DNA molecule containing the gene of interest. This is accomplished by using a specific *probe*. The probe comes from three sources, (a) from cDNA, (b) from a related organism, e.g., based on a gene sequence from a

related species. This depends on the evolutionary conservation of the sequence and often jokingly referred as "clone by phone", since the sequence can be obtained from another worker through telephone! (c) from a protein product of gene of interest – the known amino acid sequences are translated backwards to obtain the DNA sequence that encoded it.

## PCR based cloning

The PCR (Polymerase Chain Reaction) based cloning is a rapid, versatile in vivo method for amplifying a specific target DNA sequence from a pool of source DNA. In order to permit the technique to function some prior information on the nucleotide sequence of target DNA is necessary. This is because the reaction has to be primed by two short stretches (15-30 base pairs) of sequences otherwise known as oligo primers. These contain the flanking sequence information of the target DNA. They are designed in such a way that in the presence of DNA precursors (dATP, dCTP, dGTP and dTTP) and heat stable DNA polymerase enzyme, new strand will be synthesized.

The output of a PCR after 30 cycles of synthesis is about $10^5$ copies of target sequence. The PCR products, or amplicon can be visualized on an agarose gel as a discrete band of a particular size. The principal steps of a PCR reaction are (a) denaturation (b) annealing and (c) synthesis or extension. Technical simplicity and rapidity make PCR a very popular and central technology to genomic research.

## Recommended reading


Brown, T.A., 1990. *Gene cloning - An introduction*, II Edition Chapman & Hall. 286 pp.

Griffiths, A.J., J.H. Miller, D.T. Suzuki, R.C. Lewontin and W.M. Gelbart 1996. *An Introduction to Genetic Analysis*, Sixth edition W.H. Freeman and Company. 915 pp.

Old, R.W. and S.B. Primrose 1994. *Principles of Gene Manipulation – An introduction to genetic engineering*, Blackwell Scientific Publications. 474 pp.