

## **APPLICATION OF DNA FINGERPRINTING FOR THE IDENTIFICATION OF FISH PATHOGENS**

**P.C. THOMAS**

*Central Marine Fisheries Research Institute,  
PO Box 1603, Tarapuram, Kochi - 682 014, Kerala, India*

Infectious diseases cause severe economic loss to the aquaculture. The conventional diagnostic tools like culture and histological examination, immunodiffusion, counter immunoelectrophoresis, fluorescent antibody test, etc. used for detection of the pathogens are time consuming, less sensitive and not fit for mass screening. They are not effective in detecting carriers, because of which effective control of the spread of disease become difficult. The delay in diagnosis leads to heavy mortalities and huge losses from the ever increasing out breaks of infectious diseases. Therefore, the development and use of highly specific, very sensitive and rapid detection methods is the need of the day.

The DNA fingerprinting is such a technique that could be successfully used for identification and classification of the pathogens. Species/strain specific DNA fingerprint pattern of the know pathogens if worked out in advance, could be used as a frame of reference for comparing the patterns of the suspected pathogen and confirmation of its identity.

Generation of DNA fingerprint patterns can be based either on restriction fragment lengths or on polymerase chain reaction. Both of them have their own advantages and limitations.

### **DNA FINGERPRINTING BASED ON RESTRICTION FRAGMENT LENGTH**

The conventional DNA fingerprinting is based on the restriction fragment length polymorphism and consists of the restriction digestion of the DNA extracts with a panel of endonuclease enzymes individually and in various combinations. Since the sequence of the restriction site on the DNA recognised by each of these enzymes are specific, the length of each fragment resulting from the chemical cleavage of the DNA should be constant for that species, except in cases where there is a mutation of the restriction site resulting from base substitution, microdeletion/insertion at the site. The restriction fragment length patterns of the various species and strain of pathogens against a panel of enzymes if worked out and catalogued, it could be later compared with that of the unknown samples to be identified.

The procedure involves the laboratory culture of the pathogen, extraction of the DNA from it using the standard extraction protocols and restriction digestion of the purified DNA extracts with endonuclease enzyme at 37°C, running an agarose gel electrophoresis with the digested DNA sample, staining and visualization

of the bands representing the restriction fragments on a uv-transilluminator.

The inherent disadvantage of this method over the PCR assisted detection techniques is that larger volumes of DNA are needed for the analysis and this in turn necessitate culturing the pathogen isolated from the fish, prior to the extraction of DNA. Thus considerable time is involved in the procedure.

### PCR ASSISTED DETECTION OF PATHOGENS

Use of PCR assisted diagnosis has many advantages over the conventional methods. Since PCR can amplify even a single strand of DNA into millions of copies within hours, even a single bacterial cell collected from the infected fish could be used directly for the identification. Thus considerable time can be saved.

Of the various PCR assisted diagnostics, two merit special mention. They are the DNA fingerprinting based on arbitrarily primed PCR (AP PCR) and amplification of species specific virulence gene using specific primers flanking these genes.

Before going to the details of these techniques let us examine how the PCR work.

### THE POLYMERASE CHAIN REACTION

Polymerase Chain Reaction (PCR) is a technique for producing billions of copies of a specific nucleic acid sequence by performing successive rounds of *in vitro* nucleic acid replication. This is achieved by using two oligonucleotide primer that hybridise (annealing) to the opposite strands of the target DNA at positions that flank the region to be amplified and synthesis the intervening portions of both

the strands through simultaneous extension of both primers. A repetitive series of cycles involving template denaturation, primer annealing and extension of the annealed primers by DNA polymerase results in the exponential accumulation of the DNA whose termini are defined by the 5' ends of the primers. Since the primer extension products synthesised in one cycle can serve as a template for the next, the number of target DNA copies approximately doubles at every cycle. Thus 20 cycles of PCR, yields about a million fold amplification.

The tremendous applications of PCR was possible because of the availability of nucleic acid sequence information. The components of PCR reaction *viz.* template (the DNA to be amplified) primers, Taq polymerase, the deoxynucleoside triphosphate and buffer containing magnesium are assembled in a tube and the amplification reaction carried out by cycling the temperature within the reaction tube. For any given pair of oligonucleotide primers, the optimal conditions of all the above ingredients and parameters have to be standardized. Eventhough, there is no single set of conditions that will be optimal for the reaction, the conditions outlined below defines a common starting point from where modifications can be attempted.

The standard PCR mixture in addition to the sample (template) DNA contains 50 mM KCl, Tris HCl (pH 8.4), 1.5 nM MgCl<sub>2</sub>, 100 µg/ml gelatin, 0.25 µM of each primer, 200 µM of each deoxyribonucleoside triphosphate (dATP, dCTP, dGTP and dTTP) and 2.5 units of Taq polymerase. The sample DNA generally contains 10<sup>2</sup> to 10<sup>5</sup> copies of template. The volume is made upto 50 or 100 µl. The amplification is performed in a DNA thermal

cycler, each cycle consisting of denaturation at 94° C for 30 to 60 sec. annealing at 55° C for 30 to 90 sec., and extension at 72° C for 60 to 120 sec. for total of 30 cycles. Cycling could include a final extension at 72° C for 5 min. Reactions are stopped by chilling at 4° C or by addition of EDTA at 10mM.

### Materials and reagents for a PCR

#### Primers

Oligonucleotide primers in the range of 18 to 30 bases are generally used for the PCR. The sequence of the primers should be complimentary to the 3' end of the target (template) DNA strands to be amplified. Primers are the most important components that determine the success of an amplification reaction. Though there is no set of rules that will ensure the synthesis of an effective primer pair. The following guidelines are useful.

- Wherever possible, select primers with a random base distribution and with a GC content similar to that of the fragment being amplified. Avoid primers with stretches of polypurines, polypyrimidines or other unusual sequences.
- Check the primers against each other for complementarity. Use primers with low complementarity to each other. Avoid primers with 3' end overlaps in particular. This will reduce incidence of "primer dimers".

Most primers are generally 18 to 30 bases in length and the optimal length to be used in an amplification will vary. Longer primer may be synthesised, but are seldom necessary.

If shorter primers or degenerate primers are used, the thermal profile should be modified

considering the lower stability of the prime target. However, the 3' end of the primer should match the template exactly. Generally, concentrations ranging from 0.05 to 0.5 µM of each oligonucleotide should be used.

#### The reaction buffer

The components of PCR buffer, particularly the concentration of MgCl<sub>2</sub> have a profound effect on the specificity and yield of the amplification product. Concentration of about 1.5 nM is usually optimal (when 200 µM each of dNTPs are used). Excess Mg<sup>2+</sup> will result in the accumulation of non-specific amplification products and insufficient Mg<sup>2+</sup> will reduce the yield. Though several buffer formulations have been published, a consensus is beginning to emerge.

The recommended PCR buffer should contain 10mM Tris-HCl (pH 8.4) also. KCl up to 50mM can be included in the reaction mixture to facilitate primer annealing. Excess KCl inhibits Taq polymerase activity.

Gelatin or bovine serum albumin (100 µg/ml) and nonionic detergents such as Tween 20 and NP40 (0.05 - 0.1 %) are included to help stabilize the enzyme. The nonionic detergents can be replaced by 0.1% Triton X-100, but some detergent is essential.

#### Deoxynucleoside triphosphate

The deoxynucleoside triphosphate (dATP, dCTP, dGTP and dTTP) are generally used at concentrations of 200 µM each. Higher concentrations may lead to misincorporations. Low dNTP concentration reduces mispriming at non target sites. The lowest dNTP concentration appropriate for the length and composition of the target must be standardised. As a thumb rule, 20 µM of each dNTP in a 100

$\mu\text{l}$  reaction is sufficient to synthesise 10 pMol of a 400 bp sequence.

In the standard reaction, all four triphosphates were added to a final concentration of  $0.8 \mu\text{M}$ ; this leaves  $0.7 \mu\text{M}$  of the original  $1.5 \mu\text{M}$   $\text{MgCl}_2$  not complexed with dNTP. Therefore, if dNTP concentration is changed significantly, a compensatory change in  $\text{MgCl}_2$  may be necessary.

#### **Taq polymerase**

The required concentration of Taq DNA polymerase concentration is between 1 and 2.5 units per  $100 \mu\text{l}$  reaction when other parameters are optimum. When optimising a PCR, enzyme concentrations ranging from 0.5 to 5 units/ $100 \mu\text{l}$  are tested and results assayed by agarose gel electrophoresis. If the enzyme concentration is too high, non-specific background products may accumulate and if too low, an insufficient amount of desired product is made.

#### **Thermal cycle for PCR**

##### **Primer annealing**

The temperature and length of time required for primer annealing depends upon the base composition, length and concentration of the primers. As a rule of the thumb annealing temperature of  $5^\circ\text{C}$  below the true  $T_m$  of the amplification primers can be attempted. Annealing temperature in the range of  $55$  to  $72^\circ\text{C}$  generally yield the best results. At the optimal primer concentration ( $0.2 \mu\text{M}$ ) annealing will require only a few seconds.

Increasing the annealing temperature enhances discrimination against incorrectly annealed primers and reduces misextension of incorrect nucleotides at the 3' end of the primers. Therefore, stringent annealing temperature,

especially during initial few cycles will help to increase specificity.

##### **Primer extension**

Extension time depends on the length, concentration of the target sequence and temperature. Primer extensions is usually performed at  $72^\circ\text{C}$ . The rate of nucleotide incorporation at  $72^\circ\text{C}$  vary from 35 to 100 nucleotides per second depending upon the buffer, pH, salt concentration and the nature of the DNA template. An extension time of one minute at  $72^\circ\text{C}$  is considered sufficient for products up to 2 Kb in length.

##### **Denaturation**

Typical denaturation temperature is  $95^\circ\text{C}$  for 15 sec. Higher temperatures e.g.  $97^\circ\text{C}$  for 15 sec may be necessary for G+C rich targets. Denaturation steps that are too long or too high lead to unnecessary loss of enzyme activity.

##### **Cycle number**

The optimum number of cycles will depend mainly upon the starting concentration of target DNA when other parameters are optimum. Too many cycles may increase the amount and complexity of non specific background products. Too few cycles give a raw product yield.

##### **Detection and analysis of PCR product**

The PCR product will be a fragment or fragments of DNA of defined length. The simplest way to check this is to load a portion of the PCR product and molecular weight markers into an agarose gel containing ethidium bromide and carry out an electrophoresis. The DNA fragments generated by the PCR should be readily visible over an ultraviolet transilluminator.

#### **DNA fingerprinting based on arbitrarily primed PCR**

Of the different DNA techniques that can be employed for distinguishing species and strains of bacteria, DNA fingerprinting based on randomly primed PCR is the most rapid and sensitive technique. This rapid method for fingerprinting of genomic DNA described in 1990 simultaneously by Welsh & McClelland and as well as by Williams and coworkers involves the use of random primers for DNA amplification. It is a modified form of PCR. It varies from the standard PCR in the primers being used. While in the standard PCR the two primers used are those known to flank the desired portion of the target DNA, in the AP PCR method, the primers used are those not known to be flanking any particular portion of the template DNA. It therefore, tries to utilize the chance occurrence of sequences on the target DNA which are complimentary to the randomly selected primers. If the primers locate complementary sequences on the target DNA, then they anneal at the sites and under appropriate PCR conditions amplify the portion lying between two annealed sequences. The number of amplified segments depend on the number of complementary sites on the target DNA for the primer. The arbitrarily primed PCR products display differences in band patterns if the two genomic DNAs used as templates are different. This is the undelying principle in the use of arbitrarily primed PCR for fingerprinting to distinguish strains and species of pathogens. The band patterns resulting from AP PCR against an array of different primers are prepared for each species and strain of a

known pathogen to be later used as a frame of reference for comparing the patterns obtained with PCR of suspected pathogens.

One major advantage of this type of fingerprinting is that the previous knowledge of the sequence of the template DNA is not at all required. Other advantage are, the rapidity, high sensitivity and cheapness compared to other techniques like southern blotting and probes. Use of this method for identification of species and strains of bacteria viz. *Staphylococcus* and *Streptococcus* (Welsh & McClelland, 1990) *Listeria* (Mazurier & Wernars, 1992) *Borrelia* (Welsh *et al.*, 1992) and *Vibrio* (Martinez *et al.*, 1994) has been demonstrated.

#### **IDENTIFICATION OF PATHOGENS THROUGH PCR AMPLIFICATION OF SPECIES SPECIFIC GENES**

All the pathogens possess certain genes mainly related to its virulence factors or toxins which are specific to that pathogen. The core sequence of these genes are usually highly conserved. Therefore, an ideal approach for the identification of the suspected pathogen is to amplify such genes through PCR, using primers synthesised to have complementarity to the conserved regions of these genes. This implies that, previous knowledge of the sequence of the species specific gene should be available. With the increasing availability of information on the sequences of such genes of the common bacterial pathogens, application of this method shall become more common.

Given below is the list of specific genes of certain bacterial species that could be utilized for species identification.

Bacterial species	Gene
<i>Vibrio vulnificus</i>	cth (Cytotoxin haemolysin gene)
<i>Vibrio cholerae</i>	ctxA (Cholera toxin gene) ctxB
<i>V. parahaemolyticus</i> (K <sup>+</sup> )	tdh (Thermostable direct haemoysin gene)
<i>Listeria</i> <i>monocytogenes</i>	dth-18 (Delayed hypersensitivity factor) lap (Invasion associated protein gene) inIA (Internalin protein gene) hly prfA plcB ImaA
<i>Yersinia enterocolitica</i>	YadA (Virulance associated outer - membrane protein)
<i>Luminous bacteria</i>	LuxA (Luciferase gene) LuxB

The potential application of these genes for the identification of the bacterial species have been demonstrated by different workers. For example a 300 bp fragment of dth-18 gene of *Listeria monocytogenes*, amplified by Fluit *et al.*, (1993) using two 20 bp primers viz. 5'CTA ATCAAGACAATA AAA TC and 5' GTT AGT TCT ACA TCA CCT

GA was found to be useful for identifying *L.monocytogenes*. Similarly, a 564 bp fragment of the ctxA gene for *Vibrio cholerae* amplified by Fields *et al.*, (1992) using two primers of 21 bp each, was claimed to be useful in identifying *V. cholerae* bacteria.

The sensitivity and specificity of the PCR can be further enhanced manifold by carrying out nested PCR instead of a standard PCR. The process utilizes two consecutive PCRs each usually involving 25 cycles of amplification. The first PCR utilizes a pair of primes flanking the gene in question while the second PCR uses another pair of primers having complementarity to an internal segment of the gene which was amplified in the first PCR. The larger fragment produced in the first reaction is used as the template for the second PCR. The nested PCR can be 1000 times more sensitive than 50 cycles of standard PCR. Therefore, when information on the sequence species specific genes are available, amplification and visualization of that gene using a nested PCR is the method of choice when viewed from the point of sensitivity and reliability.

## REFERENCES

- Fluit, A.C., R. Torensma, M.J.C. Visser, C.J.M. Acarsman, M.J.J. Poppelier, B.H.I. Keller, P. Klapwijk and J. Verhoef 1993. Detection of *Listeria monocytogenes* in cheese with the magnetic polymerase chain reaction assay. *Applied & environmental Microbiology*, pp. 1289- 1293.
- Fields, P.I., T. Popovic, K. Wachsmuth and O. Olsvik 1992. Use of polymerase chain reaction for detection of toxigenic *Vibrio cholerae* 01 strains from the Latin American cholera epidemic. *J. Clin. Microbiol.* **30** : 2118-2121.
- Martinez, I., S. Espelied, A. Johansen, J. Welsh and M. McClelland 1994. Fast identification of species and strains of vibrio by amplification of polymorphic DNA. *J. Fish Diseases.* **17** : 297-302.
- Mazurier S.I., and K. Wernars 1992. Typing of *Listeria* strains by random amplification of polymorphic DNA. *Research in Microbiology.* **143** : 499-505.
- Welsh, J. & M. McClelland 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research.* **18** : 7213-7218.
- Welsh, J., C. Pertzman, D. Postic, I. Saint Girons, G. Baranton and M. McClelland 1992. Genomic fingerprinting by arbitrary primed polymerase chain reaction resolves *Borrelia burgdorferi* into three distinct phylectic groups. *International Journal of Systematic Bacteriology.* **42** : 370-377.