Techniques in molecular detection of seafood borne pathogens

Minimol V. A*., Sivaraman G. K. and Toms C. Joseph
Microbiology, Fermentation and Biotechnology Division
ICAR- Central Institute of Fisheries Technology, Cochin
*minimattath@gmail.com

Introduction

The conventional procedures for the detection of pathogens include selective enrichment and plating procedures for the initial screening of the pathogens that can be further identified by series of biochemical and phenotypical tests. The conventional detection and typing methods have been used for many years as a preliminary screening of pathogens surveillance and outbreaks. However, the advent in food safety practices and the increased awareness among consumers together with relatively higher occurrence of foodborne outbreaks resulted intensive investigation of the quality and safety of the products via several advanced, rapid techniques. Traditional identification of microorganisms relies on the growth of bacteria on media that are often time consuming and un-reliable whereas molecular detection assays are clearly rapid and highly specific for detection of a number of pathogens. Molecular detection methods are based on the analysis of nucleic acid so that the specificity, sensitivity and robustness of the testing protocol is much superior than the culture-based methods.

Molecular detection methods

There have been number of molecular methods have been developed for the detection of pathogens in seafood such as colony hybridization, Polymerase chain reaction methods, loop mediated amplification assays etc.

Polymerase chain reaction and its types

Polymerase chain reaction (PCR) is one of the fundamental techniques in various molecular microbiology experiments and refers to a set of procedures for the *in vitro* enzymatic amplification of a desired DNA fragment or gene from the whole genome of an organism. PCR offers the synthesis of several million copies of a target DNA sequence from a one or few copies of the sequence. PCR techniques is used widely in various diagnostics and forensic investigations, and becomes essential for many common procedures such as cloning, sequencing, microaarays etc. PCR has three main syages in which, the double stranded DNA is denatured by heat (denaturation stage) and then the temperature is lowered to allow annealing of two specific primers by complementary base pairing on the opposite strands of the DNA (annealing stage) . *Taq* polymerase directs the synthesis of the new strand from the primed sites in both directions that results in double stranded DNA (extension stage) and the procedure is repeated for 25-40 times in a thermocycler. In each cycle, the target DNA is replicated by a factor of 2 so that, after the completion of PCR, millions of copies of DNA are available for downstream applications. In addition to the amplification of a target DNA sequence by the typical PCR procedures there are several specialized types of PCR have been developed for specific applications. They are

Quantitative (Real Time) PCR

Real Time PCR is one of the PCR based assays to monitor the amplification of a particular gene /gene product in real time basis without any need for the post amplification process for visualizing DNA such as agarose gel electrophoresis, capillary electrophoresis etc. The fluorescent dye added to the reaction mixture allows the monitoring of the amplification staring from the first cycle of the PCR run and concomitantly the fluorescence is increased to 2 to 1000-fold as amplification progresses. Thus, based on the fluorescence, the DNA can be quantified over wide range of concentrations with the help of standard curves. Further, the data generated from the amplification process can easily be analyzed. The sensitivity and reliability of the result is significantly higher compared to conventional PCR. Real time PCR can be used for viral/bacterial quantification, gene/allele copy number, allelic discrimination assays (SNPs) gene expression, Methylation studies etc. The real time detection of the nucleic acid amplification is achieved by nonspecific or sequence specific strategies. The nonspecific method uses intercalating dyes which can able to produce fluorescence while binding with double stranded DNA (ds DNA). The commonly used nonspecific dye in real-time PCR is asymmetric cyanine dye called SYBR Green I. This dye has higher affinity to ds DNA compared to that of ethidium bromide and the intensity of the bound dye is higher (magnitude of 1000 folds) than the free form of syber green. This enables an increase in fluorescence during amplification. However, once the melting of the double stranded DNA after polymerisation causes the denaturation of DNA and signal strength falls off due to the detachment of fluorescent dye. Other dyes of this category include, O-PRO-1, BEBO, YOYO -1. The major advantage of nonspecific dyes are less expensive, and can be used with any pair of primers/target. The disadvantage is that it binds non-specifically to any ds DNA yielding signal from nonspecific products. However, this can be verified at the end with the help of melt curve analysis by subjecting the amplicon to a temperature range beyond its melting temperature.

The sequence specific strategies employ the use of either hydrolysis probes or hybridization probes. These probes are synthesised based on the sequences of the internal fragments of the two primers. The quantification of the PCR product is done by measuring the fluorescence signal strength based on either quenching or FRET mechanism. Hydrolysis probes are the probes which are hydrolysed due to 5'-3' exonuclease activity of DNA polymerase during the elongation stage of the PCR cycle. TaqMan Probe is widely known hydrolysis probe for RT PCR application. It is nothing but a oligo sequence labelled with reporter dye in one end (5'end) and quencher dye at the other end. In intact the fluorescence emitted from the reporter dye is banned due to the presence of quencher dye in its close proximity. During PCR run, DNA is denatured and both primer and the probe annealed to the target DNA. However, the Taq polymerase has exonuclease activity will cleave the probe and the reporter and quencher dye get separated, thus allowing the fluorescence emission from the reporter dye when it excited with a suitable light source. As amplification progress, the signal strength gets increased enabling the quantification of DNA. The melting point of the probe should be 10 degrees higher than primer Tm (melting point) as cleavage of the probe take place only during the elongation step of the PCR. In addition to TaqMan Probe, TaqMan MGB probes are also used. The Minor groove binder increase the melting temperature of the probe and it increase the duplex stability particularly for shorter probes. In case of hybridization probes the fluorescent signal is obtained due to the structure changes in the secondary structure of the

probe during hybridization phases. The changes in the structure causes increase the distance between reporter and quencher dye preventing the fluorescence resonance energy transfer (FRET) from a reporter dye to quencher dye. The probe in its intact form is a hair pin like structure and behaving non-fluorescence chromophore due to close proximity of both quencher and reporter dye. However, the conformation changes during hybridization demands separation of both dyes and the far distance among the dyes prevent the energy transfer through FRET mechanism. Thus, the increased fluorescent signal from the reporter dye enables the quantitative estimation of the DNA. With both types of assays, the exponential increase in fluorescence is used to determine the cycle threshold (Ct) which is the number of PCR cycles at which significant exponential increase in fluorescence is detected. Using a standard curve for Ct values at different DNA concentrations, quantitation of target DNA in any sample can be made.

Reverse Transcription (RT-PCR)

RT-PCR (or Reverse Transcription PCR) is used when the target nucleic acid is RNA. The central dogma in molecular biology explains about the direction or flow of information in which the DNA of the organism encodes the genetic information, intern transfer to RNA by the process of transcription and then to protein via translation process. As RNA is highly unstable and enzymatic amplification is difficult and need to reverse transcribed to cDNA for amplification. The reverse transcriptase, an enzyme that converts RNA into cDNA. This cDNA can be used for PCR and reverse transcription process may be combined in a tube, as the initial heating step of PCR being used will inactivate the transcriptase enzyme. The Tth polymerase is used for the enzymatic amplification due to its inherent RT activity, and can carry out the entire reaction. As the phenotype of an organism is explained by the RNA or protein fractions. So, RT-PCR is used in expression profiling of specific gene or gene products. It can also be used in RNA transcript analysis where in transcription start and termination sites are determined. Also, it enables the mapping of exons and introns of the gene sequence.

Nested PCR

Nested sets of primers can be used to improve PCR yield of the target DNA sequence. In nested PCR, two primer sets are used in which the first round of PCR is performed with one primer set for 15-30 cycles, then second set of primer is used for second round PCR, for an internal region of the first amplified DNA for an additional 15 to 30 cycles. The PCR product of the first round of PCR is used as DNA template for the second PCR. Thus, the nested PCR method increases the sensitivity and specificity of DNA amplification. The specificity is particularly enhanced because this technique almost always eliminates any spurious non-specific amplification products. This is because after the first round of PCR any non-specific products are unlikely to be sufficiently complementary to the nested primers to be able to serve as a template for further amplification, thus the desired target sequence is preferentially amplified. However, the increased risk of contamination is a drawback of this extreme sensitivity, and great care must be taken when performing such PCRs, particularly in a diagnostic laboratory.

Multiplex PCR

Multiplex PCR enables simultaneous amplification of many sequences or gene using two or more set of primers in one PCR. The presence of many PCR primers in a single tube could cause many problems, such as the increased formation of misprimed PCR products, "primer dimers", and the amplification discrimination of longer DNA fragments. For this type of PCR amplification, primers are chosen with similar annealing temperatures. The lengths of amplified products should be similar; large differences in the lengths of the target DNAs will favour the amplification of the shorter target over the longer one, resulting in differential yields of amplified products. In addition, Multiplex PCR buffers contain *Taq* polymerase additive, which decreases the competition among amplicon and the discrimination of longer DNA fragments during Multiplex PCR. Multiplex PCR products can be further hybridised with a gene-specific probe for verification.

Colony PCR

Colony PCR is used mainly in cloning procedure to screen the correct DNA vector constructs. Here, bacterial colonies are directly taken from the culture plate by touching a single colony using a sterile loop or tip and transferred into a PCR mix. DNA extraction from the cell is not carried out here. The denaturation step of the PCR cycle releases the DNA. In order to achieve the release of DNA from the cell, either the time period or the temperature may be extended to get an optimum amplification condition.

Loop-mediated Isothermal Amplification Assays

Loop-mediated isothermal amplification (LAMP) has been widely used to detect pathogenic bacteria in food (Zhao *et al.,* 2011). In contrast to conventional PCR, LAMP is carried out in isothermal conditions of temperature 60-65ºC with the use of specific primers. It has high DNA strand displacement activity which is mediated by *Bst polymerase* enzyme from *Geobacillus stearothermophilus*. The optimum temperature of this enzyme is 60-65ºC.The DNA strand displacement is achieved by the use of 2 sets each inner and outer primers which are specific to the target DNA. The amplification initiates with the hybridization of forward primer with the target DNA and starts the synthesis of new strand. Then, the forward outer primer hybridizes again with the same original reverse target sequence and the synthesis of this new forward strand continues until the enzyme finds the 5′ end of the first strand created with the use of the inner primer. Then owing to the property of *BSt polymerase*, the strand displacement of the first forward strand further forms a loop at one end due to the hybridization of the inner primer with target DNA. This will again serve as the template for the reverse inner and reverse outer primers and subsequently dumbbell like structure forms due to the strand displacement activity. Owing to the high displacement activity of the *Bst DNA polymerase*, a huge amount of DNA with a high molecular weight is rapidly generated. This allows target DNA amplification until 10^9 copies in less than one hour.

Molecular typing methods

Several molecular typing methods have been developed which examine the relatedness of isolates by studying their molecular composition, homology and presence or absence of specific genes etc.

Randomly amplified polymorphic DNA (RAPD)

RAPD is a typing technique based on PCR reaction in which very short nonspecific primers are used for the amplification of targeted gene (Williams *et al.*, 1990). As the primers are short, they should be able to bind many genomic sites throughout the bacterial genome. This analysis requires relatively low annealing temperature. The resulting multiple PCR products are then separated in agar gel electrophoresis. This method is simple and independent of phenotypic characters but its reproducibility from the random priming units is very low. The multiple band pattern generated by the RAPD-PCR is followed by dendrogram analysis to generate fingerprint profiles for the test organism. This method can be used to determine the clonal variations in bacterial strains. This method has been used in food borne bacterial pathogens including *V. parahaemolytcus, Escherichia coli, Salmonella, Shigella* etc.

Ribotyping

Ribotyping is rapid and specific techniques which uses the information from *rRNA* for the identification of bacteria. It involves the digestion of bacterial genomic DNA with specific restriction enzymes and the resulting fragments are separated in a gel matrix. The separated fragments are transferred to nylon membrane and hybridization will carry out with a labelled 16S or23S *rRNA* probe. Analysis of such hybridized fragments can able to identify the bacteria of interest.

Restriction Fragment Length Polymorphism (RFLP)

In PCR-RFLP, the amplified DNA is cut into short specific sequence by restriction enzymes and the resulting fragments are then separated by size using agarose gel electrophoresis. The restriction fragment profiles are very efficient in comparison of different strains. Important advantages of PCR-RFLP include inexpensiveness and lack of requirement of advanced instruments. Disadvantages include the requirements of specific restriction enzymes and difficulty to identify the variation in the nucleic acid sequence analysed. RFLP analysis has been widely used for the identification of bacterial species and biotypes.

Direct genome restriction enzyme analysis

Direct genome restriction enzyme analysis is method used for genetic diversity analysis where the DNA is cut using an endonuclease enzyme and produces a small discrete DNA fragments of 30-40 number and sizes ranging from 500-2500bp. These fragments are separated in non-denaturing polyacrylamide gel electrophoresis. Visualization of banding patterns is carried out by silver staining.

Pulsed field gel electrophoresis (PFGE)

Pulsed field gel electrophoresis (PFGE) is a typing technique widely used in epidemiological studies. It is currently recognized as a golden fingerprinting method due to the highly discriminating power as compared to other typing methods. The method involves the separation of large DNA molecules by cutting the DNA with restriction enzymes. The fragmented DNA pieces can be separated based on size using an electric field. PFGE is different from conventional DNA electrophoresis because PFGE can separate very large fragments to generate a fingerprint by constantly changing the direction of the electric field.

Analysis of fingerprint pattern is carried out by software program (BioNumerics) and that can be compared with national data base (Pulse net). PFGE is a time consuming and labor intensive method. This typing method has been used in several food poisoning outbreaks to pin point the relatedness of the strains through the space or time. This has been used in several foodborne pathogens of which the most popular is methicillin resistant *S. aureus.*

Multi locus sequence typing (MLST)

Multi locus sequence typing (MLST) is proposed in 1998 for the characterization of human pathogen *Neisseria meningitides.* Since then, it has been widely used in epidemiological and population analysis of different bacteria. This technique uses the sequences of internal fragments of usually 6 to 8 house-keeping genes or loci (Urwin and Maiden, 2003). In MLST, approximately 450-500bp internal fragments of each gene are sequenced and variations within the house keeping genes are utilized to study the genetic relatedness of the bacterial strain. An arbitrary allele number is used to denote each unique sequence of a given locus. Similarly, an arbitrary sequence type (ST) number is assigned to each unique combination of alleles (or allelic profile). Thus, it enables to identify the DNA sequence variations in a set of housekeeping genes and characterizes the strains by their unique allelic profiles and assigned sequence types.

DNA sequencing techniques

Sequencing technique provides all informations about the biochemical properties, hereditary etc. by analyzing the order of nucleic acid in polynucleotide chain of the whole DNA molecules or specific fragment or gene after amplification of the same. It has high discriminatory power, 100 % typeability and good reproducibility compared to other typing techniques. The first-generation sequencing technology includes Sanger's sequencing (Chain termination method) and chemical degradation methods. In sanger sequencing, the radioactive/fluorescent labelled deoxyribonucleotides lacking 3'hydroxy group which are unable to bind with DNA polymerase were used so that halt in the progression of extension reaction occurs thereby the resulting ddNTP bases were further run in polyacrylamide gel to yield the nucleotide sequences in the given gene fragment. In chemical degradation method, chemical cleavage of an end labelled DNA to fragments were done using specific chemicals, such a dimethyl sulphate, hydrazine etc followed by high resolution gel electrophoresis and detection by audio radiography. Short gun sequencing is one of the improvements in first generation sequencing techniques in which, the overlapping DNA fragments were cloned and sequenced separately and then assembled together to long contiguous sequence.

The second-generation sequencing includes pyrosequencing, and next generation sequencing approaches (sequence by synthesis or sequence by ligation). In pyrosequencing, the liberated pyrophosphates (two molecules of phosphate group) from each nucleotide while adding to the DNA strands during extension reactions were measured with the help of ATP sulfurylase and luciferase enzyme. The pyrophosphate and adenosine phosphosulphate reacted together in the presence of ATP sulfurylase yielded ATP and luciferin which in turn converted to fluorescent oxyluciferin compound in the presence of luciferase. Pyrosequencing approach uses this measured fluorescence from pyrophosphate synthesis. Another two approaches for the next generation sequencing include clonal amplification by bead-based emulsion PCR and

bridge PCR. Bead based PCR for sequencing is done by sequencing by ligation process in which the adapter is attached to the bead via ligation followed by water in oil emulsion PCR in which the DNA fragments gets amplified inside the droplet into millions copies. After PCR, the magnetic separation of amplified DNA beads from non-amplified DNA beads were done and sequenced by placing the beads in the sequencing slide. In bridge-based PCR, the DNA attaches to the flow cell mounted with numerous nucleotides where the DHA attaches to the complementary sequences and bend over and attached to next oilgo forming a g bridge. The polymerase enzyme synthesis the reverse strand so that the two strands releases and straighten. The result is a cluster of DNA forward and reverse strand clones. Here the sequencing is done with help of polymerase enzyme.

The third-generation sequencing techniques include single molecule real time Platform and nanopore sequencing. The PacBio Sequencing is done by passing DNA (sequence with adapter) molecule through the illuminated volume in a nano well and raw fluorescent signals from each fluorescent nucleotide when its attached to the strand during extension reaction were captured. The nucleotide sequences were determined based on the fluorescent intensities specific to each nucleotide incorporation. Nanopore sequencing technology involves the passing of DNA molecule through nanoscale pore, then the changes in electrical field surrounding the pore is measured.