MODULE 3- LECTURE 1

METHODS OF NUCLEIC ACID DETECTION

3-1.1 Introduction:

Nucleic acid molecules like deoxyribonucleic acids (DNA), ribonucleic acids (RNA) are basic, essential and primary molecules for all molecular biology related research. Before discussing the common methods to detect nucleic acids let us we learn about the isolation of pure form of nucleic acid.

3-1.2 Preparation Step:

First important step is to establish if a given sample contains DNA or RNA and whether it is in pure form, since many samples will contain both species as well as other contaminants such as cellular proteins. Irrespective of the method applied for detection, the optimal nucleic acid isolation protocol must provide

a) Reproducible result
b) No degradation of the sample and
c) Safer handling.

In deciding what method of nucleic acid measurement is appropriate, three issues are critical: specificity, sensitivity, and interfering substances. Based on these features different detection techniques are to be followed.

3-1.3 Traditional Detection Methods:

3-1.3.1 UV spectroscopy

Majority of bio-molecules intrinsically absorb light in the ultraviolet and not in the visible range. This property of UV absorbance can be used to quickly estimate the concentration and purity of DNA and RNA (also proteins) in a analytical sample. The amount of DNA in a sample can be estimated by looking at its absorbance at a wavelength of 260nm or 280nm (in the UV region). Purines and pyrimidines have absorbance maxima slightly below and above 260 respectively. Thus the absorbance maxima of different fragments of DNA vary somewhat depending on their subunit

composition. Contaminants like proteins exhibit two absorbance peaks, one between 215-230 nm (due to peptide bonds absorption) and at about 280 nm (absorption by aromatic amino acids-tyrosine, tryptophan and phenylalanine). Remember that although proteins have little absorbance at 260 nm, both proteins and nucleic acids absorb light at 280 nm. That is the reason why, if nucleic acids and proteins are mixed in the same sample, their spectra interfere (overlap) with one another.

Table: 3-1.1 The relationship between concentration of DNA, RNA, Protein and absorptivity are as below:

Sample	**Absorbance value**	***Quantity* (approximate)**
Double-stranded DNA	1at **260 nm**	**50 µg/mL**
Pure single-stranded DNA	1 at **260 nm**	**33 µg/mL**
Pure RNA	1 at **260 nm**	**40 µg/mL**
Pure protein (vary in general) Antibodies[1]	**1 at 280 nm** **1.2-1.35**	**1 mg/mL** 1 mg/mL

[1]*"Antibodies: A Laboratory Manual", by E. Harlow, and D. Lane, p. 673. Academic Press, New York., 1988.*

The purity of a solution of nucleic acid is determined by measuring the absorbance of the solution at two wavelengths, usually 260 nm and 280 nm, and calculating the ratio of A_{260}/A_{280}. Value of this ratio is 2.0, 1.8and 0.6 for pure RNA, DNA and protein respectively. A ratio of less than 1.8 signifies that the sample is contaminated with protein or phenol and the preparation is not proper.

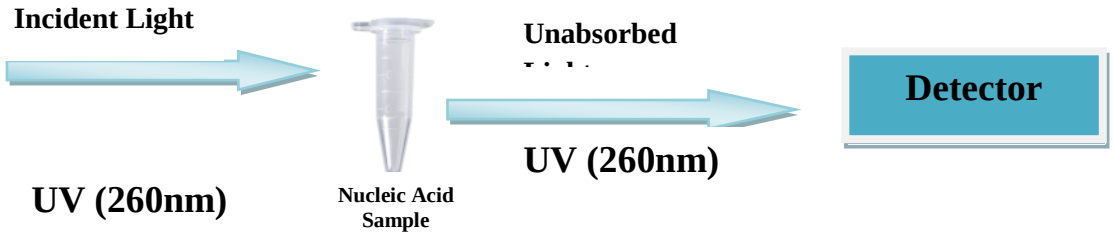


Fig 3-1.3.1 UV detection of nucleic acids

Hyperchromicity is the phenomenon of increment of absorbance when any material is exposed to UV light.The most well known application is the hyperchromicity of DNA that occurs when the DNA duplex is denatured and melting of DNA occurs. The UV absorption is increased when the duplex DNA strands are being separated into single strands, either by heat or by addition of denaturant or by increasing the pH level. On the contrary, decrease of absorbance peak is called **hypochromicity**.

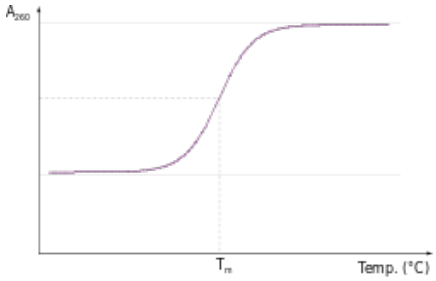


Fig. 3-1.3.1.1 -Nucleic acid melting curve showing hyperchromicity as a function of temperature

3-1.3.2 Ethidium Bromide Staining:

The IUPAC name for EtBr is 2, 7-diamino-10-ethyl-9-phenylphenanthridiniumbromide. It is commonly used as a fluorescent dye for nucleic acid staining. It binds as well as intercalates with nucleic acid (mainly with major and minor groove of DNA) and gives orange fluorescence under UV radiation from 500 – 590 nm. Usually EtBr may be added in warm agarose gel before solidification. When DNA or RNA samples are run in agarose gel electrophoresis EtBr molecules will bind with nucleic acids and help in detection under UV light. The post staining can also be done for nucleic acid detection.

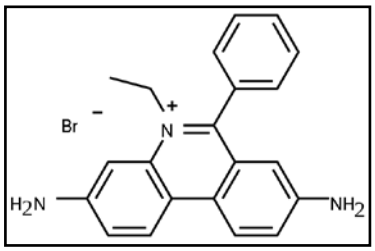


Fig 3-1.3.2 Ethidium bromide

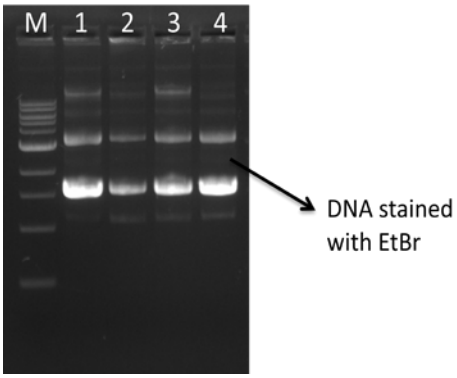


Fig. 3-1.3.2.1 Agarose gel stained with EtBr

Ethidium bromide (EtBr) is a potent mutagen and carcinogen. Dyes to stain nucleic acids such as SYBR green, SYBR Safe etc are safer to use instead of EtBr.

Polyacrylamide gel can be used for separation of Nucleic acids and post staining of the gel with EtBr is done for detection. The sensitivity of EtBr staining is of nano-molar level.

3-1.3.3 Silver Staining:

Silver staining based on reduction of silver nitrate is more sensitive than ethidium bromide for double stranded DNA, as well as detection of single stranded DNA or RNA with a good sensitivity (in picogram level). It is based on the reduction of silver cations to insoluble silver metal by nucleic acids. This chemical reaction is insensitive to the macrostructure of the DNA molecule. Reduced silver molecules deposit in the gel around the DNA bands, creating a dark black band like image (i.e. “latent image”). Then the latent image can be developed to visualize by soaking the gel in a solution of silver cations (Silver nitrate) and a reducing agent (eg. formaldehyde). The silver granules in the latent image catalyze the further reduction and deposition of silver from the solution. Bands manifest as dark brown or black regions which appear before significant background develops. Development is stopped by altering the pH of the gel to a point where silver reduction is no longer favored.

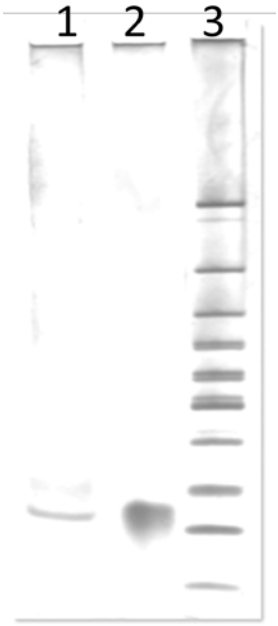


Fig. 3-1.3.3 Representation of a silver staining of DNA

Lane 1: DNA of lesser concentration
Lane 2: DNA of higher concentration
Lane 3: Low molecular weight ladder

3-1.3.4 Nanodrop:

Detection assays are persistently being developed that use progressively smaller amounts of nucleic acid, often precluding the use of conventional cuvette-based instruments for nucleic acid quantitation for those that can perform micro-volume quantitation. The patented NanoDrop microvolume sample retention system functions by combining fiber optic technology and natural surface tension properties to capture and retain small amounts of sample . This is a novel technology which allows us to measure nano-liter volumes (pico concentration) of the nucleic acid (DNA or RNA) sample. It is a type of spectrophotometer with a smaller sample size (as much less as 1-2 microlitre) requirement and higher sensitivity (even upto pico molar level). This is also a time saving technology widely used in basic molecular biology research.

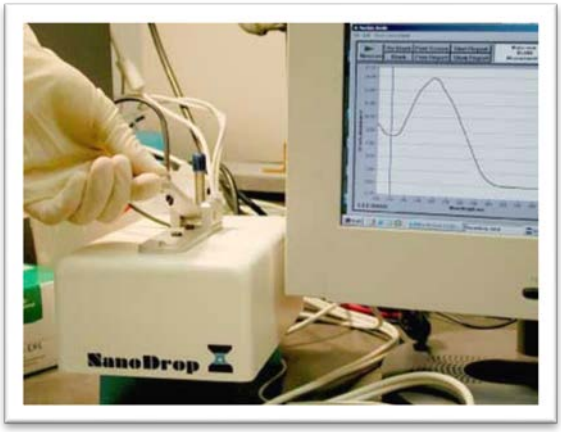


Fig. 3-1.3.4 NanoDrop® ND-1000 Spectrophotometer

(Source: http://www.biotech.wisc.edu/facilities/gec/equipment/nanodropnd1000)

3-1.3.5 Fluorometric Quantification:

Fluorometric method applies fluorescence dyes to detect the presence and concentration of a class of nucleic acid (DNA or RNA). This method is more sensitive and less prone to contaminants than UV spectroscopy.

An assay using **Hoechst 33258 dye** is specific for DNA because it is less sensitive to detect RNA. This assay is commonly used for rapid measurement of low quantities of DNA, with a detection limit of ~1 ng DNA. It is useful for the measurement of both small and large amounts of DNA (verifying DNA concentrations prior to performing

electrophoretic separations and Southern blots) because this assay accurately quantifies a broad range of DNA concentrations from10 ng/ml to15 μg/ml. The Hoechst 33258 assay can also be employed for measuring products of the polymerase chain reaction (PCR) synthesis.

Hoechst 33258 is non-intercalating reagent and binds to the minor groove of the DNA with a preference for AT sequences (Portugal and Waring, 1988). The binding to the minor groove has is dependent upon a combination of structural preferences (eg., the minor groove with a series of contiguous AT base pairs is more narrow).(Neidle (2001) ,Like other minor groove binding ligands, Hoechst 33258 is positively charged and thus form electrostatic interaction with the negative potential of stretch of AT base pairs. Upon binding to the minor groove of the double helix DNA, the fluorescence characteristics of Hoechst 33258 change dramatically, showing a large increase in emission at ~458 nm.

According to Daxhelet *et al*, the fluorochrome 4',6-diamidino-2-phenylindole (DAPI) has similar characteristics to H33258 and binds to the minor groove as well. DAPI is also appropriate for DNA or RNA quantitation, although it is not as commonly used as Hoechst 33258. DAPI is excited with a peak at 344 nm. Emission is detected at around 466 nm for DNA, similar to Hoechst 33258 but for RNA the peak shifts to ~500 nm.

3-1.3.6 Hybridization-Based Techniques

Hybridization-Based Techniques for nucleic acid detection has higher resolution (down to the actual nucleotide sequence) and utilizes "probe sequence" for DNA or RNA, and when it finds its intended target, binds to it by hybridization process. Since the "probe" is attached to a label such as a fluorescent chemical isotope, the bound sequence can thus be visualized. This is the principle of **Southern Blotting** method, as well as its variants, and requires the target nucleic acid sequence to undergo separation by agarose-electrophoresis, transferred to an appropriate membrane (typically nitrocellulose), and then treated with a solution containing the labeled probe (fluorescent or colorimetric).

Northern blotting involves the use of electrophoresis to separate RNA samples by size and detection with a hybridization probe complementary to part of or the entire target sequence. Since the probe specifically binds to its target, the membrane can be documented and analyzed so that the bound target sequences can be identified and studied. The detailed methodology is described in later chapter (Module-4Lecture 3).

FISH (Fluoroscence in-situ hybridization) for visualization of nucleic acids developed as an alternative to older methods that used radio labeled probes (Gall and Pardue, 1969). Several drawbacks of isotopic, radiolabeled hybridization stimulated the development of novel techniques like FISH. In the year 1980, RNA was first directly labeled on the 3' end with fluorophore was used as a probe for specific DNA sequences (Bauman et al., 1980). Enzymatic incorporation of fluorophore-modified bases throughout the probe has been widely used for the preparation of fluorescent probes. The use of amino-allyl modified bases (Langer et al., 1981), which could later be conjugated to any sort of hapten or fluorophores, was critical for the development of in situ technologies because it allowed production of an array of low-noise probes by simple chemistry. Methods of indirect detection result into higher magnitude signal output by the use of secondary reporter molecules that bind to the hybridization probes.

In the early 1980s, assays like nick-translated, biotinylated probes, and secondary detection

by streptavidin conjugates were used for detection of DNA (Manuelidis *et al*., 1982) and mRNA (Singer and Ward, 1982) targets. Later, advanced labeling of synthetic, single-stranded DNA probes allowed the chemical preparation of hybridization probes carrying enough

fluorescent molecules to allow direct detection (Kislauskis *et al*., 1993). At this current era, based on these reaction themes of indirect and direct labeling have since been introduced, a wide spectrum of nucleic acid detection schemes are available.

Whereas the initial development of FISH involved expansion of the types of probe and number of detectable targets, the outlook for future development of fluorescence techniques will include expansion of the subjects of investigation like in clinical, diagnostics, forensic applications.

Other detection methods include amplification of target region by polymerase chain reaction and various forms of chemiluminescent, fluorescent or radioactive detection method. Specific technical or sample requirements are present for each method, and the original purpose for detection the nucleic acid will determine which of the methods is most suitable.

Bibliography:

Bauman, J. G., Wiegant, J., Borst, P. and van Duijn, P. (1980). A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochrome labelled RNA. Exp. Cell Res. 128, 485-490.

Daxhelet, G.A., Coene, M.M., Hoet, P.P., and Cocito, C.G. 1989. Spectrofluorometry of dyes with DNAs of different base composition and conformation. Anal. Biochem. 179:401-403.

Fluorometric Quantitation of DNA Using Hoechst 33258 Cold Spring HarbProtoc; 2006; doi:10.1101/pdb.prot4458.

Gall, J. G. and Pardue, M. L. (1969). Formation and detection of RNA-DNA hybrid molecules in cytological preparations. Proc. Natl. Acad. Sci. USA 63, 378-383.

Kislauskis, E. H., Li, Z., Singer, R. H. and Taneja, K. L. (1993). Isoform specific 3'untranslated sequences sort alpha-cardiac and beta-cytoplasmic actin messenger RNAs to different cytoplasmic compartments. J. Cell Biol. 123, 165-172.

Langer, P. R., Waldrop, A. A. and Ward, D. C. (1981). Enzymatic synthesis of biotin-labeled polynucleotides: novel nucleic acid affinity probes. Proc. Natl. Acad. Sci. USA 78, 6633-6637.

Manuelidis, L., Langer-Safer, P. R. and Ward, D. C. (1982). High-resolution mapping of satellite DNA using biotin-labeled DNA probes. J. Cell Biol. 95, 619-625.

Molecular Cloning- A laboratory manual-David W. Russell & Joseph Sambrook.

Neidle, S. DNA minor-groove recognition by small molecules. 2001. *Nat. Prod. Rep.* 18:291-309.

Portugal, J. and Waring, M.J. (1988). Assignment of DNA binding sites for 4,6-diamidine-2-phenylindole and bisbenzimide (Hoechst 33258): A comparative footprinting study. *Biochem. Biophys. Acta* 949:158-168.

Singer, R. H. and Ward, D. C. (1982). Actin gene expression visualized in chicken muscle tissue culture by using in situ hybridization with a biotinated nucleotide analog. Proc. Natl. Acad. Sci. USA 79, 7331-7335.

MODULE 3- LECTURE 2

POLYMERASE CHAIN REACTION (PCR) AND ITS APPLICATIONS

3-2.1 Introduction:

Polymerase chain reaction (PCR) is a widely employed technique in molecular biology to amplify single or a few copies of DNA, generating millions of copies of a particular DNA sequence. The polymerase chain reaction results in the selective amplification of a target region of a DNA or RNA molecule. PCR has been extensively exploited in cloning, target detection, sequencing etc. The method consists of thermal cycles of repeated heating followed by cooling of the reaction mixture to achieve melting and primer hybridization to enable enzymatic replication of the DNA.

3-2.2 History:

By 1971, a “repair synthesis" process was reported which was an artificial system containing primers and templates that can allow DNA polymerase to copy target gene. The DNA polymerases initially employed for *in vitro* experiments were unable to withstand these high temperatures. In 1976, Chien *et al* discovered a novel DNA polymerase from the extreme thermophile *Thermus aquaticus* which naturally dwell in hot water spring (122 to 176 °F). The enzyme was named as *Taq* DNA polymerase which is stable upto 95°C. In 1985, Kary Mullis invented a process Polymerase Chain Reaction (PCR) using the thermo-stable *Taq* polymerase for which he was awarded Nobel Prize in 1993.

Fig. 3-2.2 PCR Thermo cycler

(*Adapted from http://products.invitrogen.com/ivgn/product/4452300***)**

3-2.3 Basic Protocol for Polymerase Chain Reaction:

3-2.3.1 Components and reagents:

A basic PCR set up requires the following essential components and reagents :

1. Template DNA containing the DNA region (target) to be amplified.

2. Primers that are complementary to the 5' ends of each of the sense (Forward primer) and anti-sense strand of the DNA target (Reverse primer).

3. *Taq* polymerase or other thermostable, high fidelity DNA polymerase (Pfu polymerase isolated from *Pyrococcus furiosus*).

4. Deoxyribonucleotide triphosphates (dNTPs), which are the building-blocks for a newly synthesized DNA strand.

5. Buffer solutions to provide a suitable chemical condition for optimum activity and stability of the DNA polymerases.

6. Divalent cations (eg. magnesium or manganese ions). They act as a co-factor for Taq polymerase which increases its polymerase activity. Generally Mg^{2+} is used, but Mn^{2+} can be applied to achieve PCR-mediated DNA mutagenesis. This is because higher Mn^{2+} concentration leads to higher error rate during DNA synthesis.

3-2.3.2 Procedure:

Typically, PCR is designed of 20-40 repeated thermal cycles, with each cycle consisting of 3 discrete temperature steps: denaturation, annealing and extension. The thermal cycles are often proceeded by a temperature at a high range (>90°C), and followed by final product extension or brief storage at 4 degree celsius. In PCR cycles, the temperatures and the duration of each cycle is determined based on various parameters like the type of DNA polymerase used, the melting temperature (Tm) of the primers, concentration of divalent ions and dNTPs in the reaction etc. The various steps involved are:-

a) Initial Denaturation

b) Denaturation

c) Annealing

d) Extension

e) Final extension

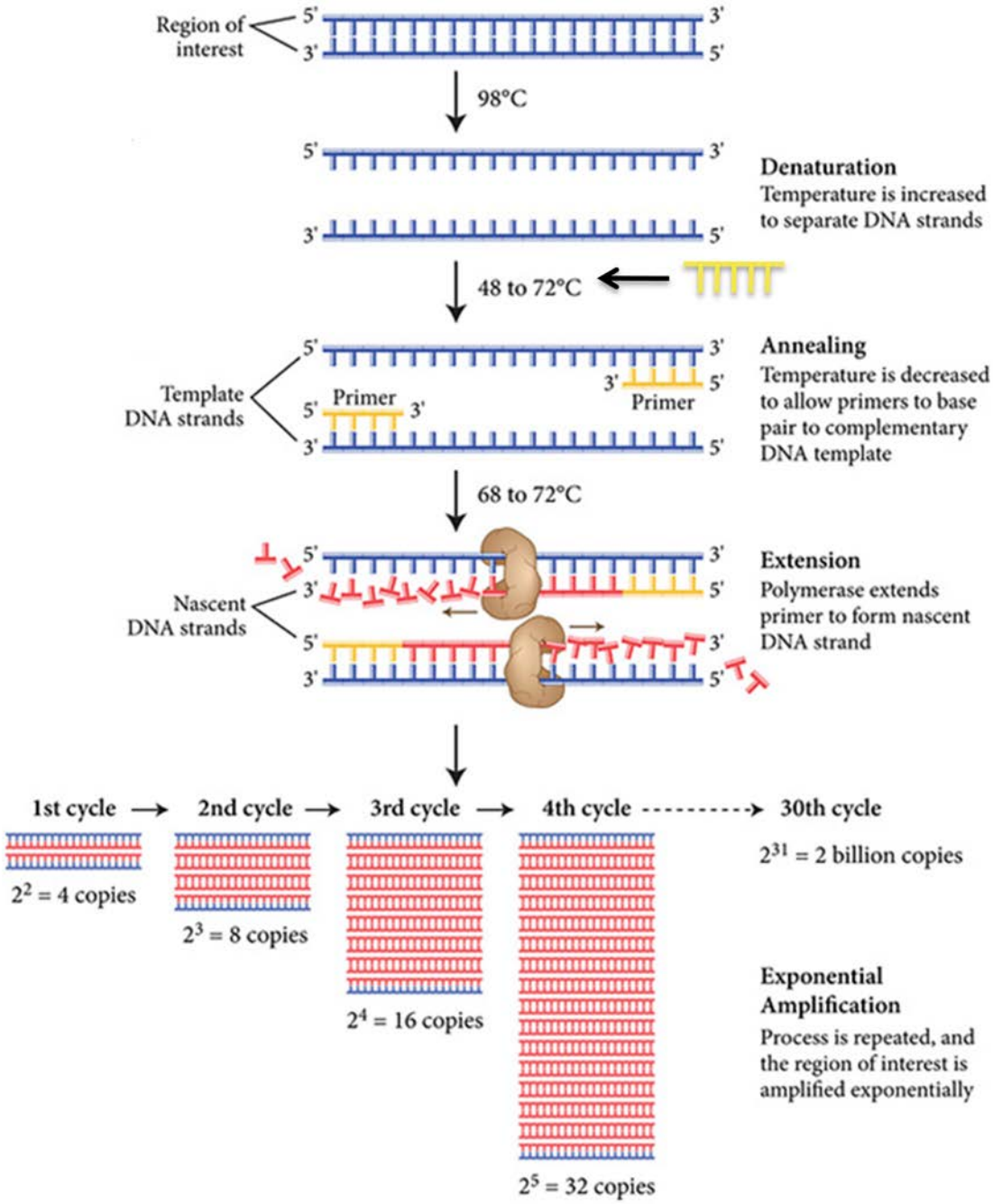


Fig 3-2.3.2 The sequential steps of PCR

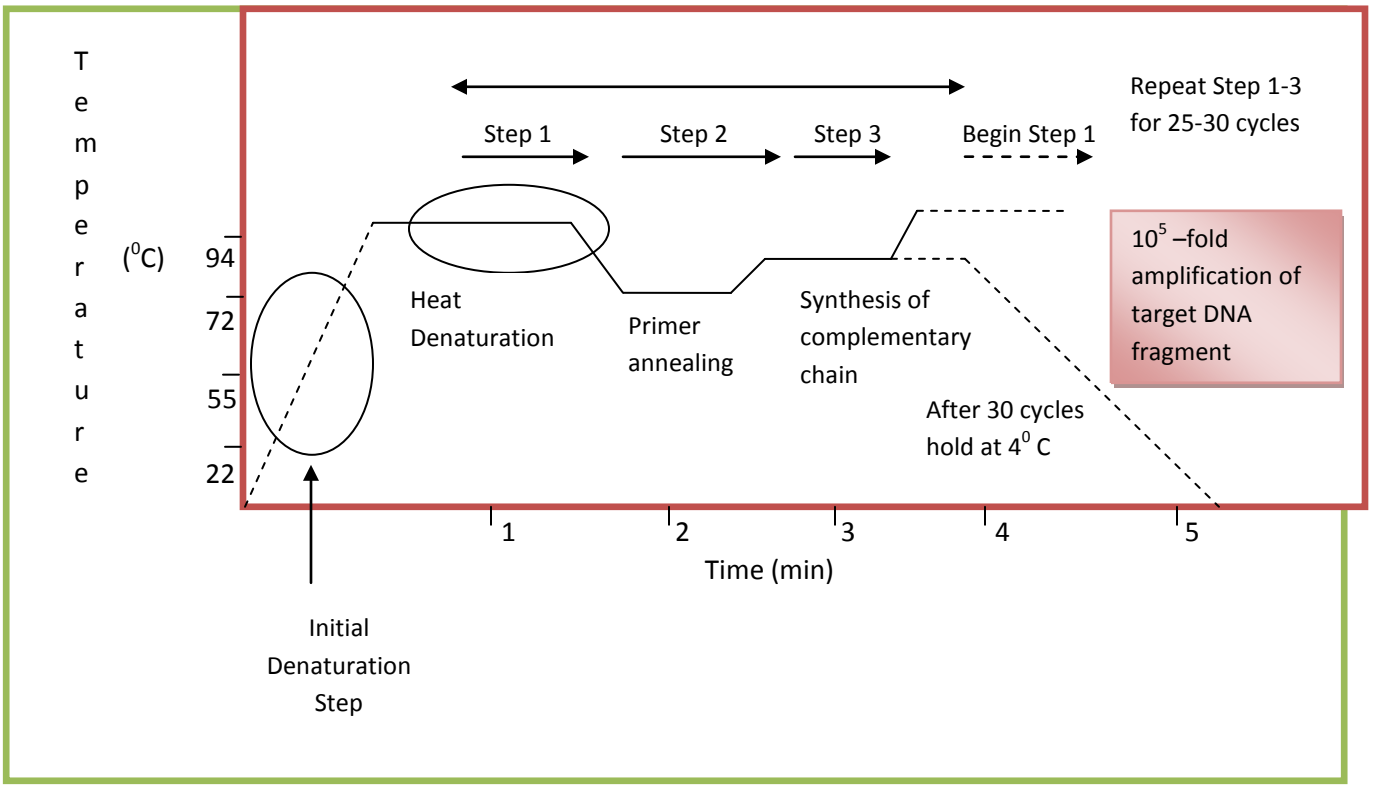


Fig 3-2.3.2.1 Basic Thermal Profile of PCR

3-2.3.2.1 Initial denaturation:

Initial denaturation involves heating of the reaction to a temperature of 94–96 °C for 7-10 minutes (or 98 °C if extremely thermostable polymerases are used). For specifically engineered DNA polymerases (Hot start Taq polymerases) activity requires higher range of temperature. The initial heating for such a long duration also helps in gradual and proper unfolding of the genomic DNA and subsequent denaturation, and thus exposing target DNA sequence to the corresponding primers.

3-2.3.2.2 Denaturation:

Denaturation requires heating the reaction mixture to 94–98 °C for 20–30 seconds. It results in melting of the DNA template by disrupting the hydrogen bonds between complementary bases, yielding single-stranded DNA molecules.

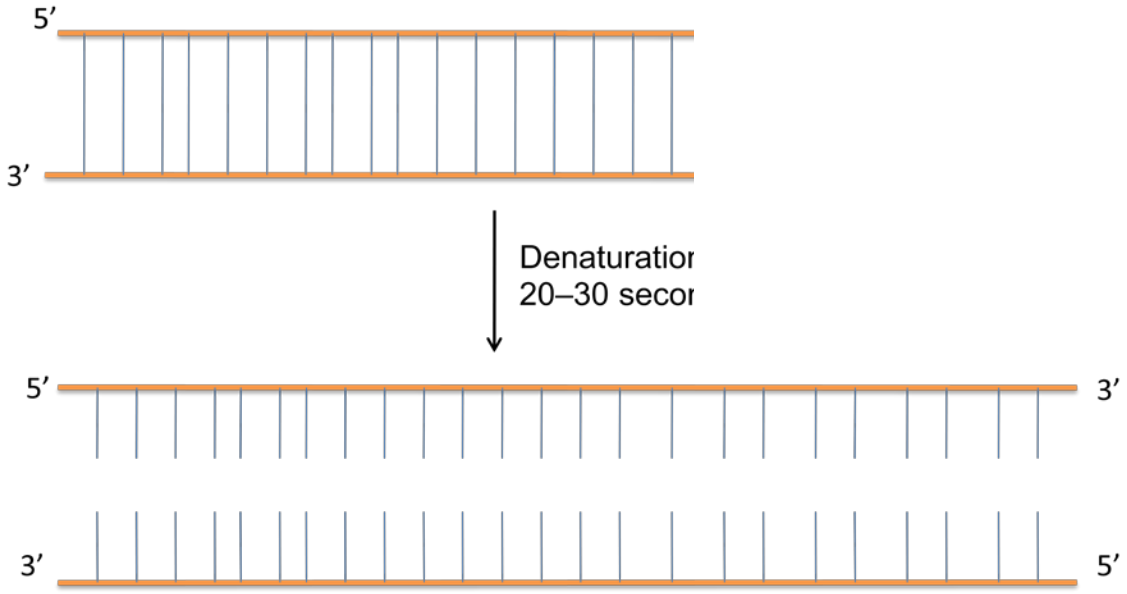


Fig3-2.3.2.2: Denaturation of double stranded DNA to single stranded DNA

3-2.3.2.3 Annealing:

Following the separation of the two strands of DNA during denaturation, the temperature of the reaction mix is lowered to 50–65 °C for 20–50 seconds to allow annealing of the primers to the single-stranded DNA templates. Typically the annealing temperature should be about 3-5 °C below the T_m of the primers. Stable complimentary binding are only formed between the primer sequence and the template when there is a high sequence complimentarity between them. The polymerase enzymes initiate the replication from 3' end of the primer towards the 5'end of it.

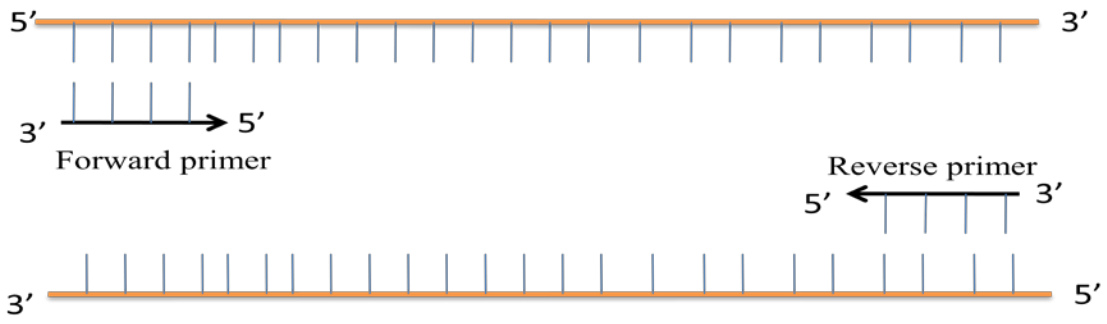


Fig3-2.3.2.3: Annealing of primer

3-2.3.2.4 Extension/Elongation:

Extension/elongation step includes addition of dNTPs to the 3’ end of primer with the help of DNA polymerase enzyme. The type of DNA polymerase applied in the reaction determines the optimum extension temperature at this step. DNA polymerase synthesizes a new DNA strand complementary to its template strand by addition of dNTPs, condensing the 5'-phosphate group of the dNTPs with the 3'-hydroxyl group at the end of the nascent (extending) DNA strand. Conventionally, at its optimum temperature, DNA polymerase can add up to a thousand bases per minute. The amount of DNA target is exponentially amplified under the optimum condition of elongation step. The drawback of Taq polymerase is its relatively low replication fidelity. It lacks a 3' to 5' exonuclease proofreading activity, and has an error rate measured at about 1 in 9,000 nucleotides.

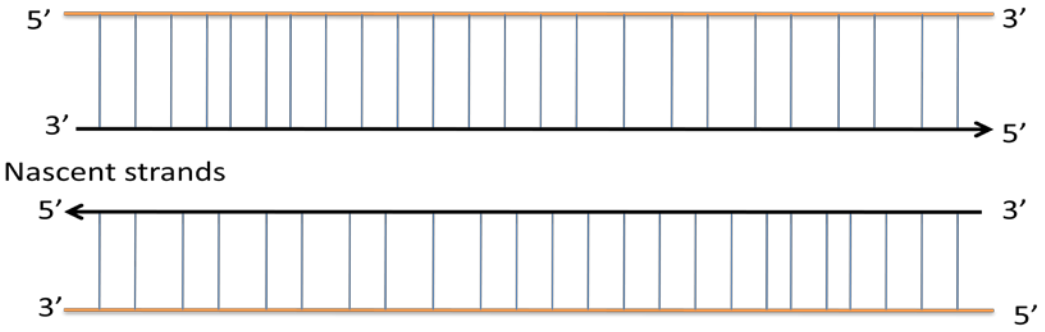


Fig3-2.3.2.4: Extension/ Elongation

3-2.3.2.5 Final elongation & Hold:

Final elongation step is occasionally performed for 5–15 minutes at a temperature of 70–74 °C after the last PCR cycle to ensure amplification of any remaining single-stranded DNA.

Final hold step at 4 °C may be done for short-term storage of the reaction mixture.

After around 30 cycles of denaturation, annealing and extension, there will be over a billion fragments that contain only your target sequence. This will yield a solution of nearly pure target sequence.To check the desired PCR amplification of the target DNA fragment (also sometimes referred to as the amplicon or amplimer), agarose gel electrophoresis is employed for separation of the PCR products based on size. The

determination of size(s) of PCR products is performed by comparing with a DNA ladder, which contains DNA fragments of known size, run on the gel along side the PCR products.

3-2.4 Key factor affecting the polymerase chain reaction: Designing of Primers

The specificity of the PCR depends crucially upon the primers. The following factors are important in choosing effective primers.

a) Primers should be 17 to 30 nucleotides in length. However for certain studies the RAPD primer length might be of 10-12 nucleotides.

b) Ideally GC content of the primers is 50% .For primers with a low GC content, it is desirable to choose a long primer so as to avoid a low melting temperature.

c) The basic formula to calculate melting temperature is

T_m = 4°C x (number of G's and C's in the primer) + 2°C x (number of A's and T's in the primer)

Two primers must have a similar Tm value. In case of several primer candidates, we have to choose primers which have the higher Tm value among them.

d) Sequences with long runs (i.e. more than three or four) of a single nucleotide should be avoided.

e) Primers with significant secondary structures (self –hair pin, loop formation) are undesirable.

f) There should be no base complementarities between the two primers.

3-2.5 Applications

3-2.5 .1 Infectious disease diagnosis, progression and response to therapy

PCR technology facilitates the detection of DNA or RNA of pathogenic organisms and, as such, helps in clinical diagnostic tests for a range of infectious agents like viruses, bacteria,protozoa etc. These PCR-based tests have numerous advantages over conventional antibody-based diagnostic methods that determine the body's immune response to a pathogen. In particular, PCR-based tests are competent to detect the

presence of pathogenic agents in-advance than serologically-based methods, as patients can take weeks to develop antibodies against an contagious agent. PCR-based testshave been developed to enumerate the amount of virus in a person's blood (‘viral load') thereby allowing physicians to check their patients' disease progression and response to therapy.This has incredible potential for improving the clinical management of diseases caused by viral infection, including AIDS and hepatitis, assessment of viral load throughout and after therapy.

PCR-based diagnostics tests are available for detecting and/or quantifying a number of pathogens, including:

1.HIV-1, which causes AIDS

2.Hepatitis B and Cviruses, might lead to liver cancer

3.Human Papillomavirus,might cause cervical cancer

4.*Chlamydia trachomatis*, might lead to infertility in women

5.*Neisseria gonorrhoeae*, might lead to pelvic inflammatory disease in women

6.Cytomegalovirus, might cause life threatening disease in transplant patients and other immunocompromised people, including HIV-1/AIDS patients

7.*Mycobacterium tuberculosis*, which in its active state causes tuberculosis and can lead to tissue damage of infected organs.

3-2.5 .2 Diagnosis of genetic diseases

The use of PCR in diagnosing genetic diseases, whether due to innate genetic changes or as a result of a natural genetic mutations, is becoming more common. Abnormality can be diagnosed even prior to birth. **Single-strand conformation polymorphism** (**SSCP**), or single-strand *chain* polymorphism, is defined as conformational difference of single-stranded nucleotide sequences of identical length as induced by differences in the sequences under certain experimental conditions. These days, SSCP is most applicable as a diagnostic tool in molecular biology. It can be used in genotyping to detect homozygous individuals of different allelic states, as well as heterozygous individuals who inherit genetic aberrations.

3-2.5 .3 Genetic counselling

Genetic counselling is done for the parents to check the account of genetic disease beforehand to make a decision on having children. This is of course governed by national laws and guidelines. Detection of genetic disease before implantation of an embryo in IVF (*In vitro* fertilisation) also known as pre-implantation diagnosis can also be done exploiting PCR based method. Further to diagnose inherited or a spontaneous disease, either symptomatic or asymptomatic (because of family history like Duchene muscular dystrophy) PCR based method is very useful.

3-2.5.4 Forensic sciences

Genetic fingerprint is one of the most exploited application of PCR (also known as DNA profiling).Profiles of specific stretches of DNA are used in genetic fingerprinting (generally 13 loci are compared) which is differ from person to person. PCR also plays a role in analysis of genomic or mitochondrial DNA, in which investigators used samples from hair shafts and bones when other samples are not accessible.

3-2.5 .5 Research in Molecular Biology

PCR is an essential technique in cloning procedure which allows generation of large amounts of pure DNA from tiny amount of template strand and further study of a particular gene. Some alterations to the PCR protocol can generate mutations (general or site-directed) in a sequence either by an inserted fragment or base alteration.PCR is used for sequence-tagged sites (STSs) as an indicator that a particular segment of a genome is present in a particular clone. A common application of Real-time PCR is the study of expression patterns of genes during different developmental stages. PCR can also investigate ‘ON or OFF” of particular genes at different stages in tissues(or even in individual cells).

3-2.5.6 Others

PCR has numerous applications in various fields. The Human Genome Project (HGP) for determining the sequence of the 3 billion base pairs in the human genome, relied heavily on PCR.The genes associated with a variety of diseases have been identified using PCR. For example, Duchenne muscular dystrophy,which is caused by the mutation of a gene, identified by a PCR technique called Multiplex PCR. PCR can help to study for DNA from various organisms such as viruses or bacteria. PCR has been used to identify and to explore relationships among species in the field of evolutionary biology. In anthropology, it is also used to understand the ancient human migration patterns. In archaeology, it has been used to spot the ancient human race. PCR commonly used by Paleontologists to amplify DNA from extinct species or cryopreserved fossils of millions years and thus can be further studied to elucidate on.

Bibliography:

Bruce A. White (1993); Methods in molecular biology: PCR Protocols, *Springer*, Vol 15**.**

Julian Burke (1996) PCR: Essential techniques; John Wiley & Sons.

Module 3: Lecture 3

VARIATIONS IN PCR AND THEIR APPLICATIONS

3-3.1 Introduction:

In the last lecture we discussed that in 1985 Kary Mullis developed Polymerase chain reaction (PCR) which has become an everyday laboratory procedure to obtain millions of copies of a specific piece of DNA of interest. The reaction requires the template DNA, target sequence specific primers and mixture of dNTPs, and heat-stable *Taq* DNA polymerase. A typical PCR involves repeated cycles of heating and cooling of the reaction mixture to allow template DNA denaturation, primer annealing and new DNA strand formation. Billions of amplicons will be generated at the end of 30-35 repetitive PCRcycles. At the end of 30 cycles, the specific sequence will be exponentially amplified to generate multiple copies (2^{30} copies) (amplicons).The basic PCR process has been modified to expand its applications in various fields. PCR has been a widely used technique as a diagnostic and research tool. The applications of PCR are progressively growing and are used in many scientific disciplines including genetics, molecular biology, forensic science, paternity testing, clinical diagnostics, microbiology, environmental science, hereditary studies etc. Depending upon the application there may be a variation in the PCR technique deployed which is another advantage of this method.

3-3.2 Variations in PCR:

3-3.2.1 Reverse Transcription PCR (RT-PCR):

In reverse transcription polymerase chain reaction (RT-PCR), first a RNA strand (template) is reverse transcribed into its complementary DNA copy using reverse transcriptase, and subsequently cDNA is amplified using PCR.Various types of *Reverse transcriptase* enzyme, isolated from *Avian myeloblastosis virus (AMV),Moloney murine leukemia virus (MMLV* or *MuLV)* are generally used to produce a DNA copy from RNA template. Random primers, an oligo (dT) primer or sequence-specific primers are used to amplify cDNA. Alternatively, some thermostable DNA polymerases (e.g., *Tth* DNA

polymerase isolated from*Thermus thermophilus*) having reverse transcriptase activity, which requires manganese (Mn^{2+})as a cofactor for activation instead of magnesium. Basic PCR follows this initial reverse transcription step for amplification of the target sequence.

RT-PCR is widely used in the diagnosis of genetic disorders and semi quantitatively in the calculation of specific expression level of particular RNA molecules within a cell or tissue. RT-PCR also helps in obtaining eukaryotic exon sequences from mature mRNAs.

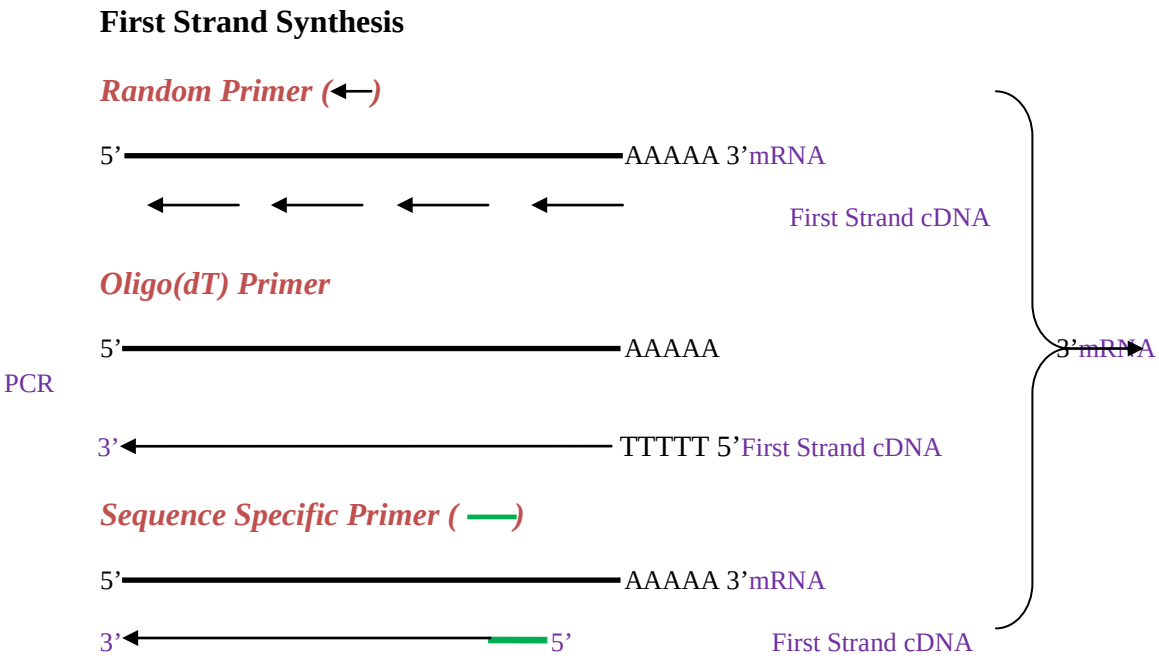


Fig 3-3.2.1 Schematic Diagram of RT PCR

3-3.2.2 Real Time PCR or Quantitative PCR (qPCR):

Real-time or Quantitative PCR (qPCR) uses the linearity of DNA amplification to quantify absolute or relative amounts of target sequence in a sample. With the help of a fluorescent reporter, the amount of generated DNA can be measured. In qPCR, DNA amplification is monitored at each cycle of PCR. When the DNA is getting amplified logarithmically at each cycle, the amount of fluorescenceincreases over the basal level. The thermal cycle at which the signal exceeds the fluorescence detection threshold is

known as the **Threshold cycle**(C_T) or crossing point. A standard curve of log concentration against C_T can be made by making use of multiple dilutions of a known amount of standard DNA. The quantity of DNA or cDNA in an unknown sample can thus be determined from its C_T value.

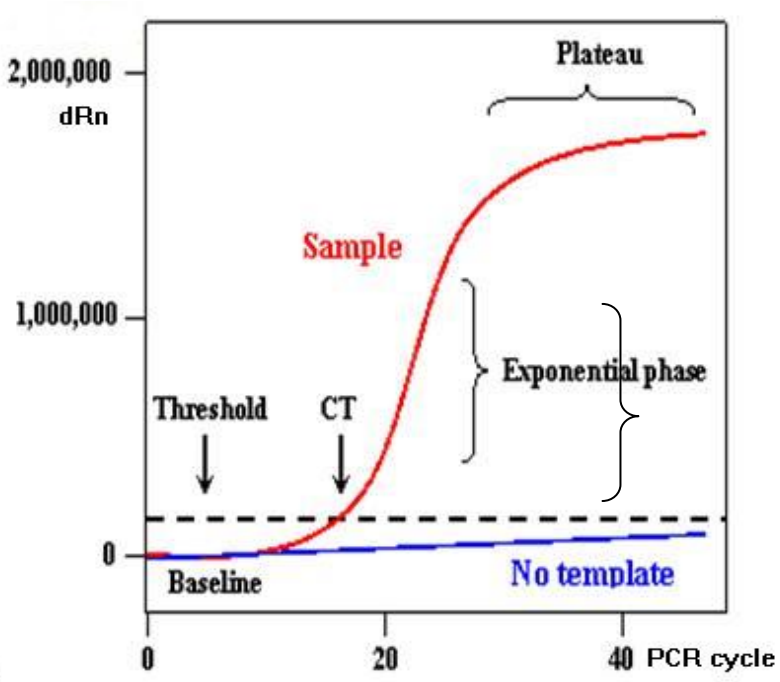


Fig 3-3.2.2 A Representation of Real Time PCR Plot (Obtained from NCBI Probe)

*(**Baseline** is defined as PCR cycles in which a reporter fluorescent signal is accumulating but is beneath the limits of detection of the instrument. **dRn** is an increment of fluorescent signal at each time point.)*

There are two detection chemistries used for q PCR:

a) Use of an intercalating dye like the SYBR® Green I dye which incorporates between the base pairs of DNA. This detection method is suitable when the PCR reaction generates a specific product, as the dye is capable of intercalating into any double-stranded DNA product.

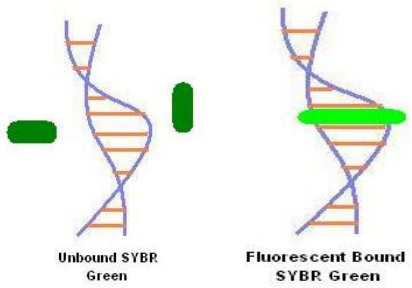


Fig 3-3.2.2.1 SYBR Green I Detection Mechanism

(The dye intercalates with double-stranded DNA in the reaction. In the bound state, SYBR Green I exhibits 1000 fold more fluorescence than the unbound state. The fluorescence signal increases in proportionwith the increase in amplified DNA)

b) Use of primer or short oligonucleotide specific to the target of interest, as in TaqMan® probes, Molecular Beacons™, or Scorpion primers. In case of molecular Beacons, they are labeled with a fluorescent dye or quencher and do not exhibit any significant fluorescence in the free, unhybridized condition. But upon binding to the template, the probe becomes fluorescent as the quencher gets distanced from the fluorescent reporter. The amount of PCR product amplified is directly proportional to the amount of fluorescence.

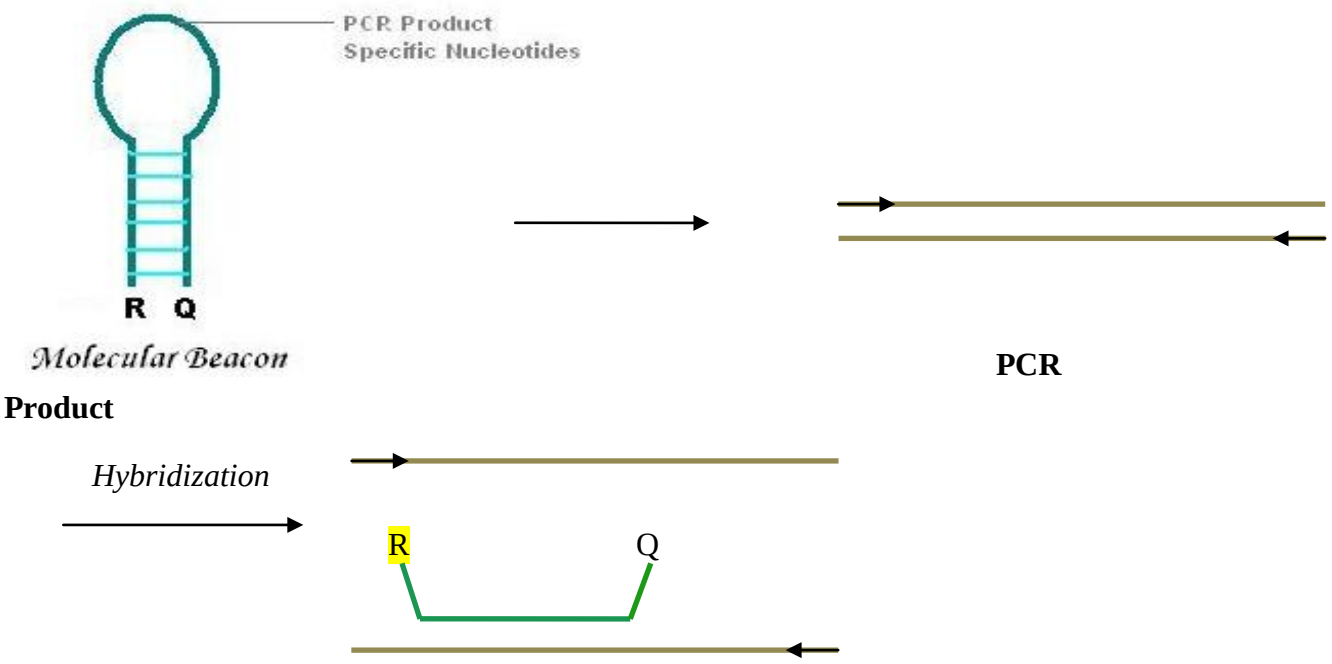


Fig 3-3.2.2.2 Detection of PCR Product by Molecular Beacon

While in the case of TaqMan® probes, fluorescence occurs when the dye is clipped from the probe during the polymerase extension.

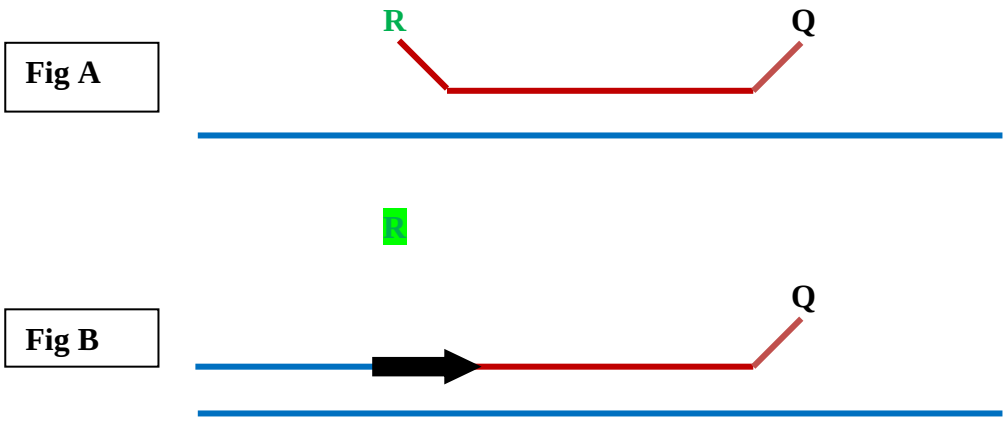


Fig 3-3.2.2.3 explains detection of qPCR product by TaqMan Probe.

[A] In TaqMan probe, when Reporter is adjacent to Quencher –there is no fluorescence.

[B] During extension, polymerase hydrolyzes the probe leading to separation of Reporter from the Quencher and thus fluorescence is exhibited.

Scorpion primers are similar to Molecular Beacons™, but they contain a PCR primer sequence. When the target DNA is amplified during PCR, the beacon fragment binds to the newly synthesized DNA, and thus, separates the fluorophore from the quencher. To prevent the stem-loop structure from being copied during PCR, a "PCR blocker" is also incorporated in the hairpin. The intra-molecular binding between the scorpion primer and the downstream PCR product is kinetically more favorable than that of Molecular Beacons and TaqMan probes. Unlike the TaqMan and Molecular Beacons, the Scorpion system do not require a separate probe and hence, gives a larger fluorescent signal.

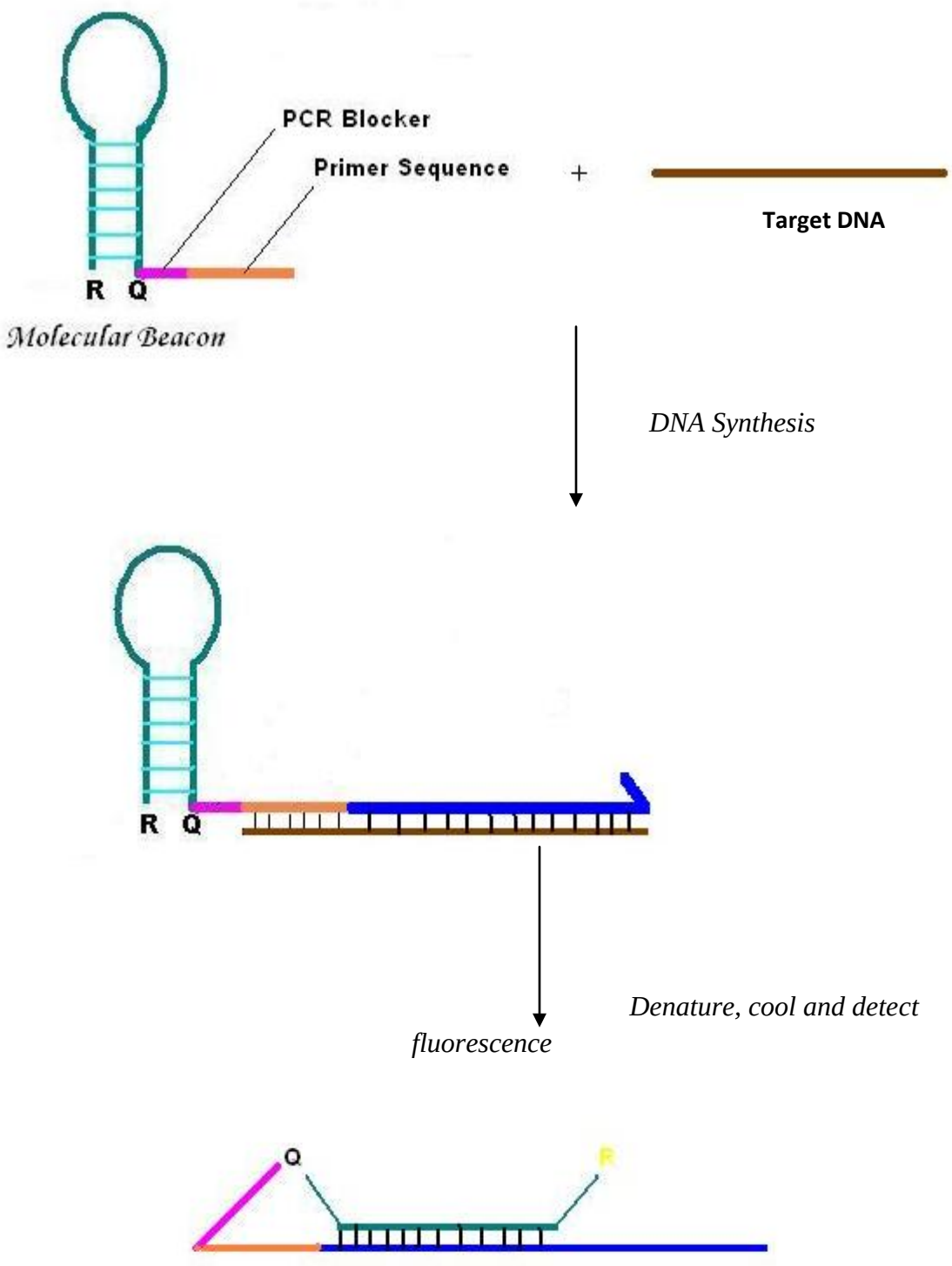


Fig 3-3.2.2.4 Scorpion Primer Detection Mechanism

3-3.2.3 Hot Start PCR:

During conventional PCR, the *Taq*DNA polymerase remains active at room temperature butits activity falls when temperature is lowered down. In some occasions, nonspecific primer binds at low temperatures. This nonspecific bound primer can then be extended by the *Taq*DNA polymerase, resulting in nonspecific products generation and lowering of product yield. Hot Start PCR helps to mitigate nonspecific primer annealing, primer dimer formation, and often enhances product yield. The method involves withholding one of the critical components of PCR like magnesium ion, the enzyme DNA polymerase, PCR primers, and dNTPs from the reaction until the temperature in the first cycle rises above the annealing temperature. It can be performed manually by heating the reaction components to the melting temperature(Tm) of DNA (e.g. 95°C) before adding the polymerase. Another strategy involves the use of a physical barrier like wax to segregate key reaction components. Specialized reagents provide an alternative route by inhibiting the polymerase activity at room temperature, either by the binding of an antibody, or by covalently bound inhibitors that is required for dissociation a high-temperature activation step. ‘Hot-start/cold-finish PCR' is mediated by genetically engineered polymerases that are inactive at ambient temperature but is activated at high temperatures.

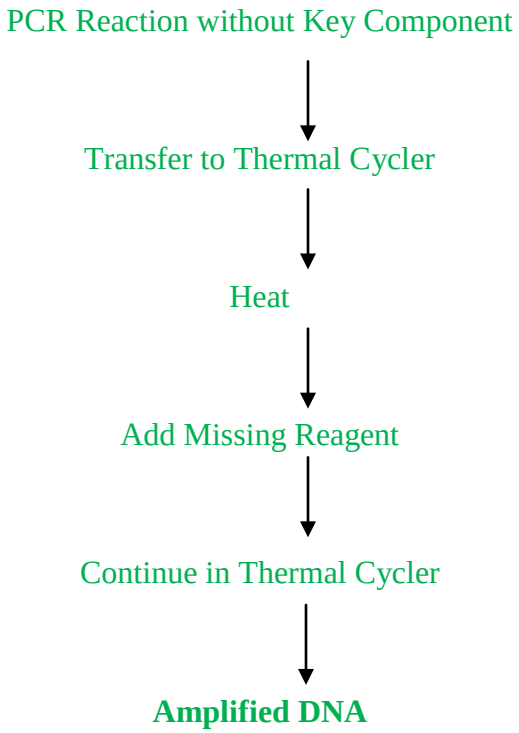


Fig 3-3.2.3 Sequential Steps of Hot-Start PCR

3-3.2.4 Colony PCR:

In Colony PCR, bacterial colonies are directly used for PCR amplification and screening for example, screening for positive transformed DNA vector constructs. Colony samples are obtained with a sterile pipette tip or toothpick and transferred into a PCR mix. For cell disruption, the PCR is either run with an extended time at 95°C (when standard polymerase is used), or with a reduced denaturation step at 100°C (when special recombinant DNA polymerase is used).It is a widely used technique for rapid screening of transformed colonies to check the presence of insert, its orientation and the size. It is a simpler method as steps involving genomic or plasmid DNA isolation, restriction digestion and Southern blotting are not required at all. It can also be used for screening desired recombinant clones from DNA libraries, thus reducing the time and effort required for screening of large number of colonies.

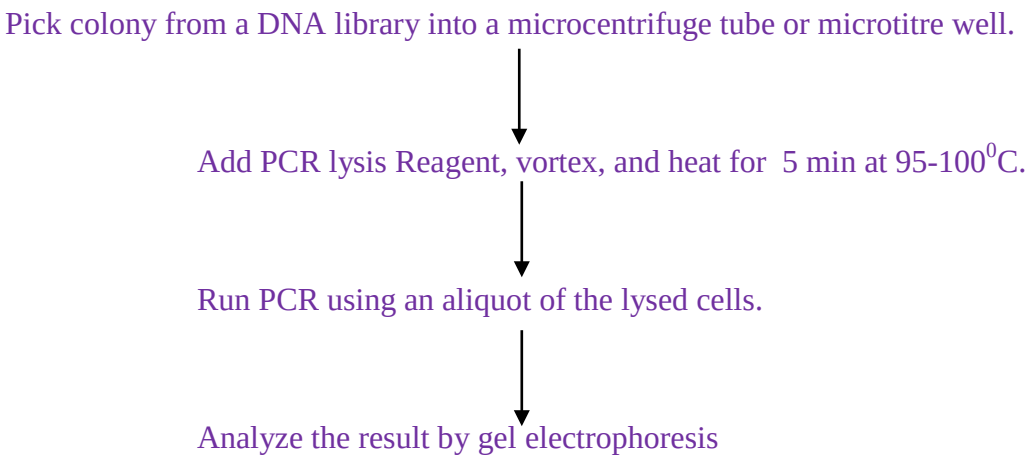


Fig3-3.2.4: Sequential Steps of Colony PCR

3-3.2.5 Nested PCR:

Nested PCR is a modified form of PCR designed to reduce the contamination in products occurs due to non-specific amplification or non-specific primer binding sites. Nested PCR makes use of two sets of amplification primers. The target DNA sequence for one set of primers (termed **"inner"** primers) is situated within the target sequence for the second set of primers (termed **"outer"** primers).In practice, "outer primers" are first used in a standard PCR procedure for a test sample. Then "inner primers" are used in a second PCR reaction where the product generated in the first PCR serves as the amplification target for the second PCR reaction. In this procedure, the sensitivity of the assay is increased to multiple fold as the product of the first reaction is re-amplified in the second reaction. The specificity of the assay is improved as the inner primers will only amplify if a specific product is obtained in the first PCR reaction.

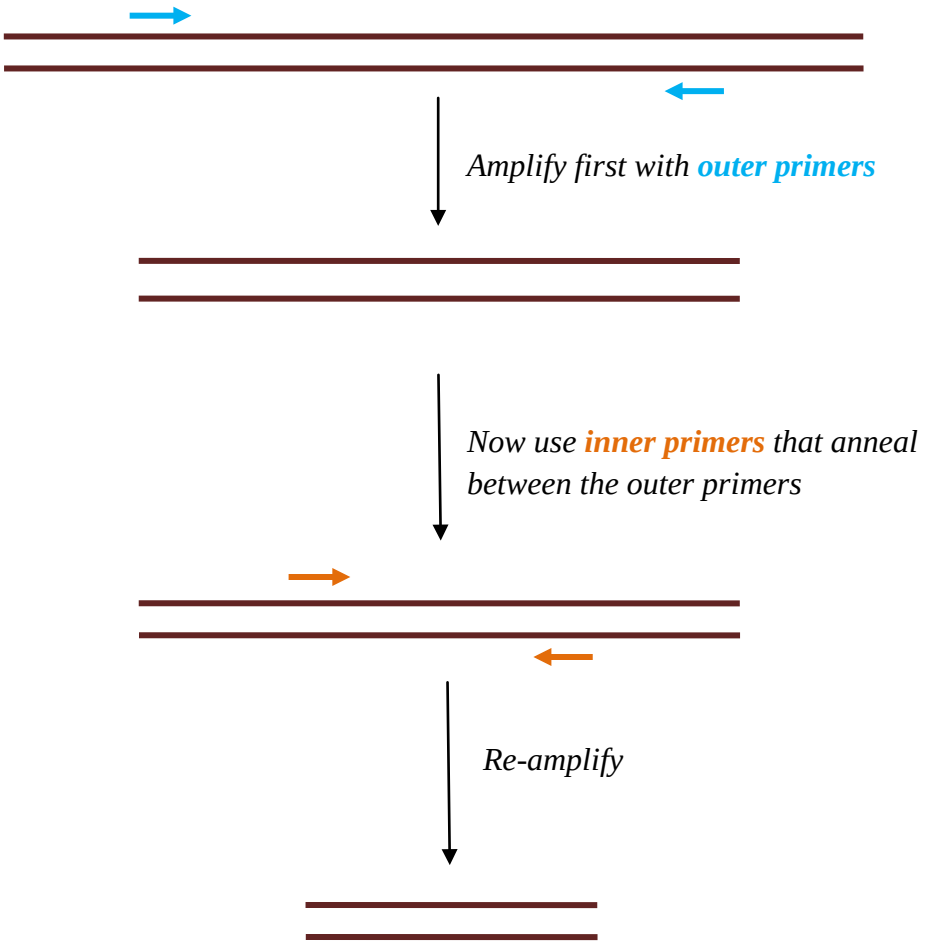


Fig3-3.2.5: Schematic Diagram of Nested PCR

3-3.2.6 Touchdown PCR:

Touchdown (TD) PCR offers a simple and rapid means to optimize PCR specificity, sensitivity and yield, instead of lengthy optimizations and primer redesigning.TD-PCR uses an initial annealing temperature above the estimated melting temperature (T_m) of the primers in use, then progressively moves down to lower, more tolerant annealing temperature as the PCR cycle continues successively. Any difference in melting temperature between the correct and incorrect annealing will result in an exponential twofold advantage per cycle. TD-PCR has special application for amplification of templates that are usually difficult to amplify. It can also be used as a standard to enhance specificity and product formation.TD-PCR has widespread applicability in standard PCR protocols like reverse transcriptase PCR, in the construction of cDNA libraries and in the development of screens for detecting Single nucleotide polymorphisms.

3-3.2.7 INVERSE PCR:

The standard PCR suffers from the limitation that the 5' and 3' flanking regions of target DNA must be known. Inverse PCR methodology helps to overcome this drawback by making feasible amplification by a PCR reaction when only one internal sequence is known. It follows standard PCR program, but the primers are oriented in the reverse direction with respect to the normal orientation. A restriction fragment that self-ligates to form a circle serves as the template in this reaction. It involves:

- Restriction endonuclease mediated digestion of target DNA.
- Induction of self-ligation to form a circular product.
- It is then restriction digested with a known endonuclease to generate a cut within the known internal sequence.
- The resultant product with known terminal sequences can now be used for standard PCR.

Inverse PCR has many applications in molecular biology like the identification of genomic inserts and the amplification and identification of sequences flanking transposable elements.

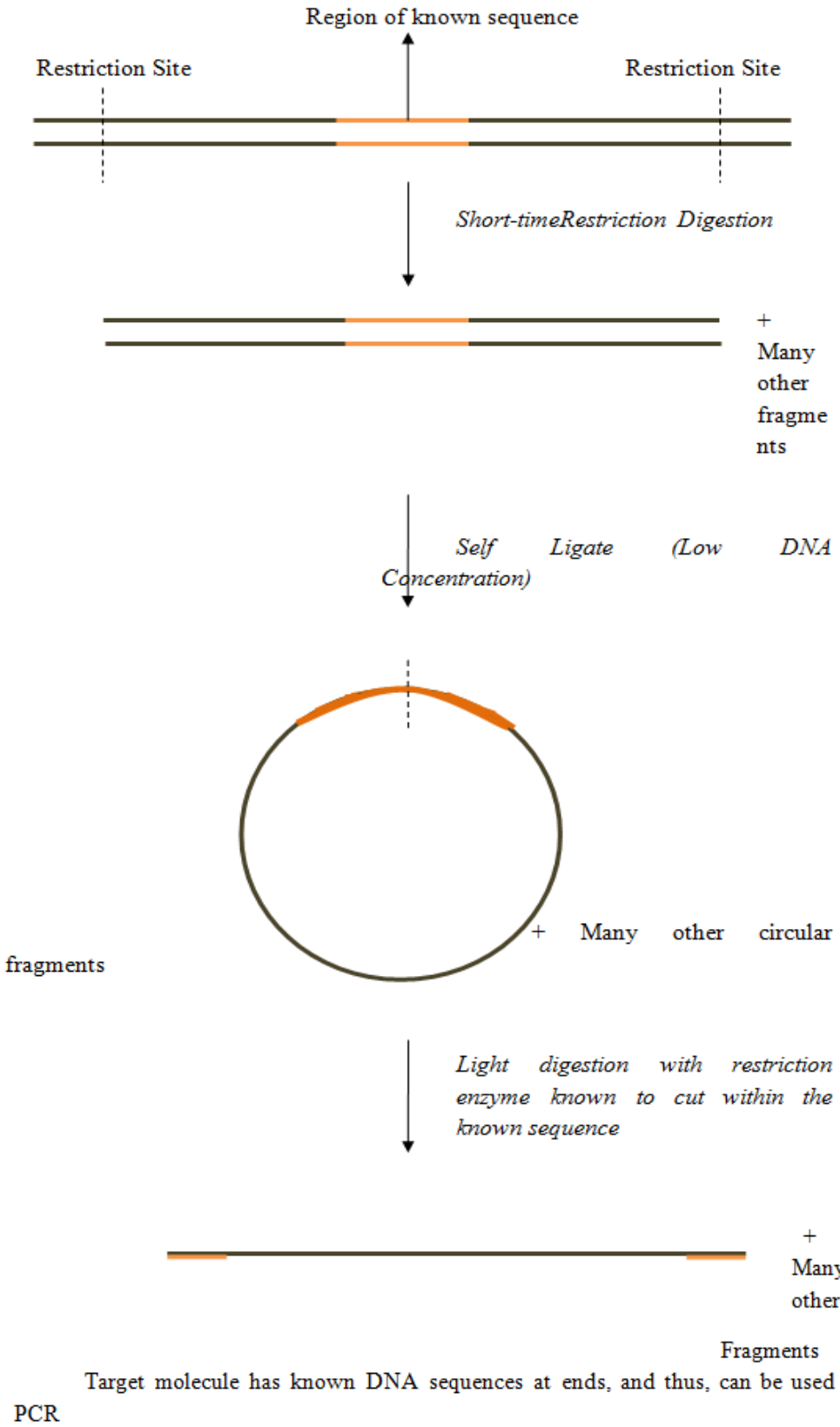


Fig 3-3.2.7: Schematic Diagram of Inverse PCR

3-3.2.8 *In-situ* or Slide PCR

In-situ PCR or slide PCR has been used since the 1990s to perform PCR directly on small tissue samples, tissue micro arrays, or other small cell samples, rather than extracting DNA or RNA from the samples. In typical *in-situ* PCR protocol, first PCR master-mix directly applied onto the sample (mounted on a slide) and then mixture of sample and reagent are covered with a cover slip. The slide is then applied on a regular thermo cycler equipped with an *in-situ* adaptor or slide adaptor. Initially, as a result of the stoichiometry of nucleic acid components, the reaction becomes primer-driven. As the reaction proceeds, the generated PCR product becomes a target for further amplification and promotes further elongation. The sequences that have been elongated in multiple reactions are accumulated in the fixed cell, which serve as a target in the detection phase for the binding of radiolabeled probe. Another version of in situ PCR, known as reverse transcriptase *in situ* PCR (RT-*in situ*-PCR) has even been used to detect RNAs and thus analyzing the expression of gene of interest.

3-3.2.9 Multiplex PCR:

In multiplex PCR assay, more than one target sequence can be amplified by using multiple primer pairs in a single reaction mixture. Since multiple genes are the target in a single reaction, various information comes out from a single test run otherwise it would require several reactions. However, the primer sets must have annealing temperatures within a narrow range and are optimized to work properly within a single reaction.

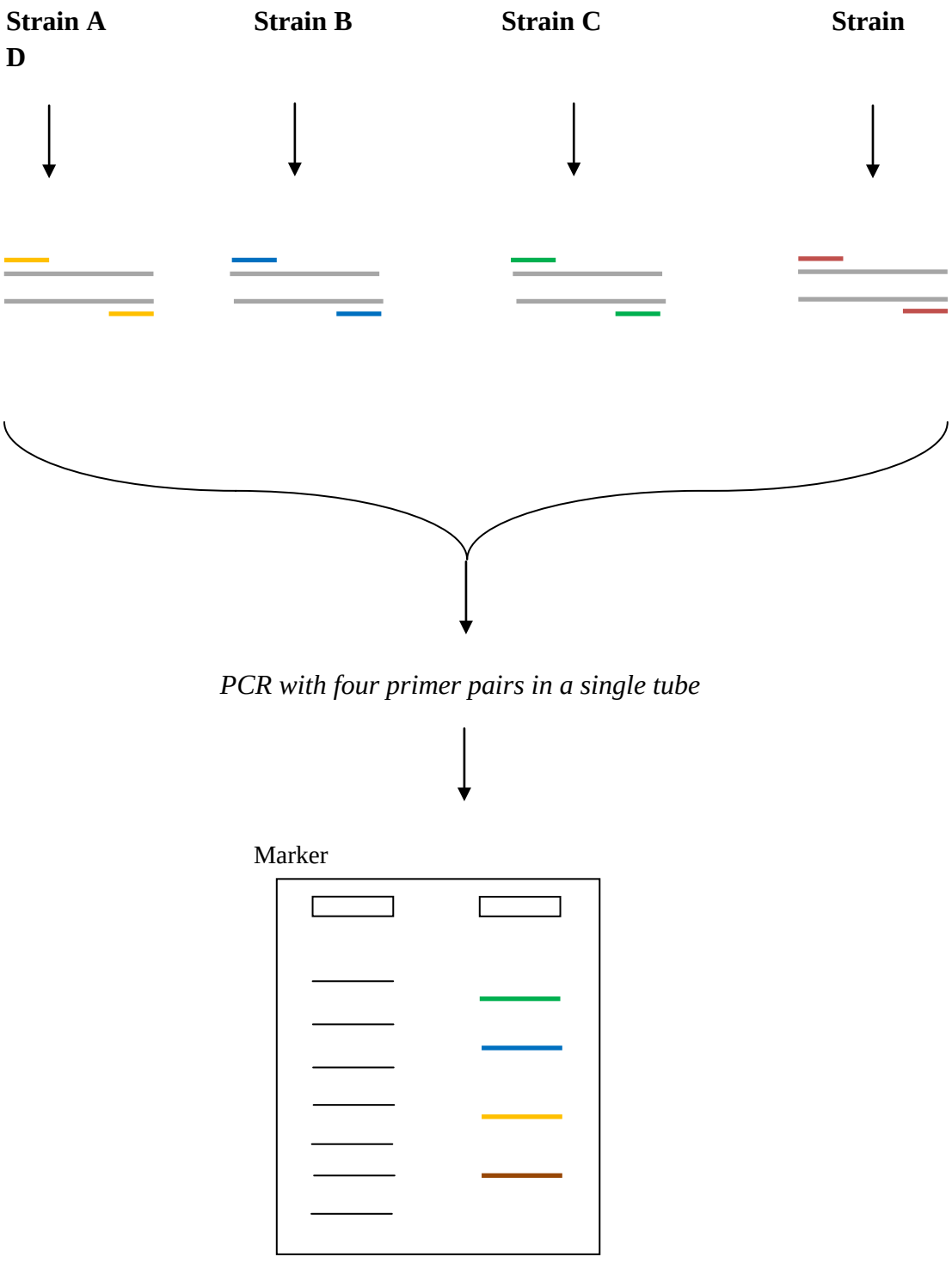


Fig3-3.2.9: Schematic Diagram of Multiplex PCR

In multiplex assays, false positive results are often obtained because each amplified fragment provides an internal control for the other amplified fragments. It is ideal for conserving costly polymerase and templates in short supply. Multiplex PCR can be utilized to determine the amount of a particular template in a sample by exponential amplification and internal standards. It is widely used in:

a) Pathogen identification,

b) High throughput SNP genotyping,

c) Mutation analysis,

d) Gene deletion Analysis,

e) Template quantitation,

f) Linkage analysis,

g) RNA detection,

h) Forensic studies etc.

3-3.2.10 Assembly PCR:

Assembly PCR, also known as Polymerase cycling assembly, is a method for the generation of large DNA oligonucleotides from shorter fragments. This process also requires DNA hybridization and annealing along with DNA polymerase to amplify a complete DNA sequence in a precise order based on the single stranded oligonucleotides used in the process. The reaction mixture contains oligonucleotides ~50 base pairs long each overlapping by about 20 base pairs. The reaction with all the reagents is then carried out for ~30 cycles and by an additional 23 cycles with the help of end primers. It thus facilitates construction of the synthetic genes and even entire synthetic genomes.

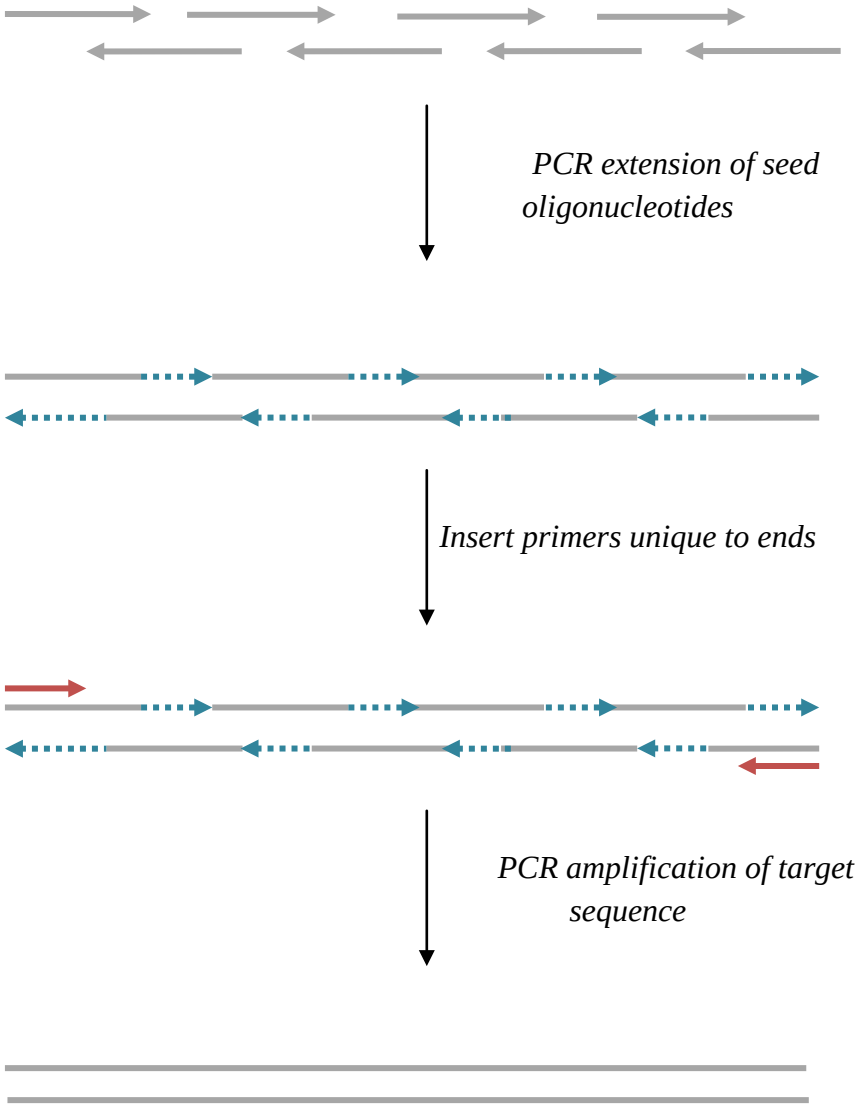


Fig3-3.2.10: Schematic Diagram of Assembly PCR

3-3.2.11 Solid Phase PCR:

Solid phase PCR uses surface-bound primers instead of freely-diffusing ones to amplify DNA. The primers are bound to the glass surface via a 5′-specific linkage that can withstand PCR conditions and leaves the 3′-ends available for DNA polymerase activity. The newly synthesized DNA can be detected and quantified by radioactive and fluorescent hybridization assays. It offers a probable approach for attachment of DNA molecules by their 5′-end on a solid support and can be used as an alternative route for producing DNA chips for genomic studies.

3-3.2.12 Asymmetric PCR:

In asymmetric PCR preferential amplification of one strand of the target DNA occurs. Here, one of the primers has a limiting concentration or is left out. When the limiting primer has exhausted, arithmetic increase in replication occurs through elongation of the excess primer. It is used to generate one DNA strand as product for use in sequencing methods and in probing hybridization.

3-3.2.13 LATE PCR:

Linear-After-The-Exponential-PCR (or LATE-PCR) is a modification of asymmetric PCR. In usual asymmetric PCR, the limiting primer concentration decreases mid-reaction and thus decreases the reaction efficiency. Instead, in LATE PCR, a limiting primer with a higher melting temperature (Tm) in comparison to the excess primer is used, and there by allows asymmetric synthesis of strand of interests in excess while maintaining the reaction efficiency.

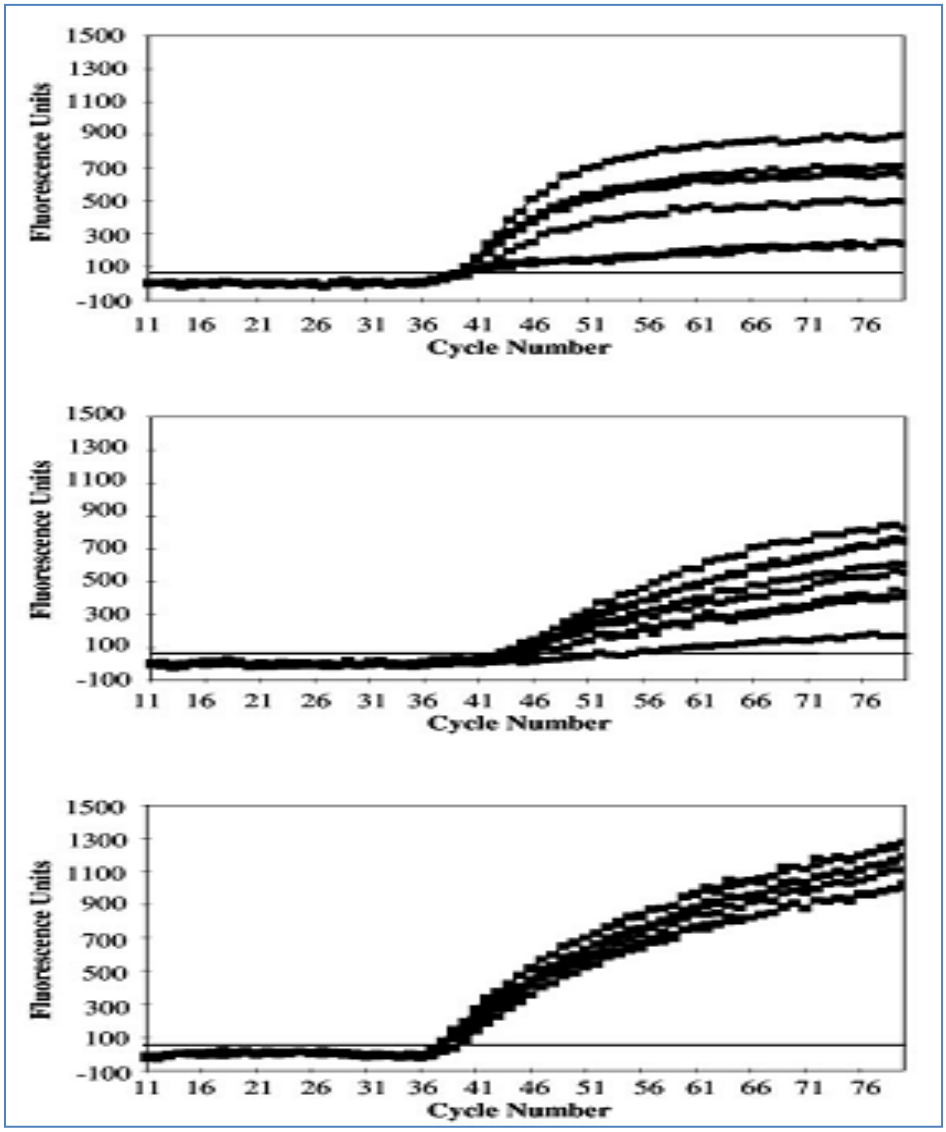


Fig 3-3.2.13 Comparison of symmetric PCR, standard asymmetric PCR, and LATE-PCR

(respectively from top to bottom,* Reference: Sanchez et al 2004)***

Symmetric PCR reaches a plateau around 50 cycles and shows considerable variability in kinetics with a simultaneous large end-point range. While, conventional asymmetric PCR is flawed by delayed linear kinetics and lower end-point signals. Thus, LATE-PCR only shows the efficiency in linear kinetics for over than 80 cycles.

Bibliography:

Innis MA, Myambo KB, Gelfand DH, Brow MA (1988). "DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA". Proc. Natl. Acad. Sci. U.S.A.85 (24): 9436–40. doi:10.1073/pnas.85.24.9436.

Korbie DJ, Mattick JS (2008) Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat Protoc*3: 1452–1456.

Paul, N., Shum, J., Le, T. Hot Start PCR in Methods in Molecular Biology in RT-PCR *Protocols, Springer Science&Business Media*, LLC 2010, 630: 301-18.

Pavlov AR, Pavlova NV, Kozyavkin SA, Slesarev AI (2006). "Thermostable DNA Polymerases for a Wide Spectrum of Applications: Comparison of a Robust Hybrid TopoTaq to other enzymes". In Kieleczawa J. DNA Sequencing II: Optimizing Preparation and Cleanup. Jones and Bartlett. pp. 241–257. ISBN 0-7637-3383-0.

Sanchez JA, Pierce KE, Rice JE, Wangh LJ. Linear-after-the-exponential (LATE)-PCR: an advanced method of asymmetric PCR and its uses in quantitative real-time analysis. *Proc Nat AcadSci* USA 2004; 101:1933– 8.

Sandy B. Primrose & Robert W. Old; Principles of Gene Manipulation, 7th Edition.

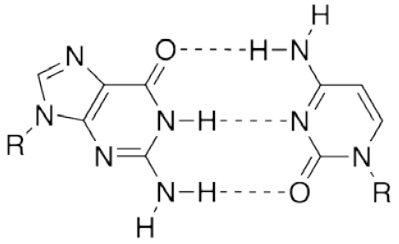
Singh OP, GoswamiGeeta, Nanda N, Raghavendra K, Chandra D, and Subbarao SK. 2004. “An allele-specific polymerase chain reaction assay for the differentiation of members of the *Anopheles culicifacies* complex.”

Module 3 Lecture 4

METHODS OF NUCLEIC ACID HYBRIDIZATION

3-4.1 Introduction:

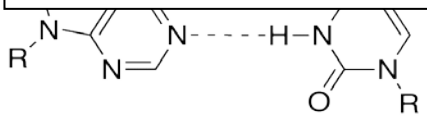
Nucleic acid hybridization is a basic technique in molecular biology which takes advantage of the ability of individual single-stranded nucleic acid molecules to form double-stranded molecules. According to Watson-Crick base pairing, adenine binds with thymine and guanine binds with cytosine by hydrogen bonding.



Guanine Cytosine

A GC base pair demonstrating three intermolecular hydrogen bonds

An AT base pair demonstrating two intermolecular hydrogen bonds



Adenine Thymine

3-4.2 Southern hybridization:

The basic principle behind the southern hybridization is the nucleic acid hybridization. Southern hybridization commonly known as southern blot is a technique employed for detection of a specific DNA sequence in DNA samples that are complementary to a given RNA or DNA sequence. It was the first blotting technique to be devised, named after its pioneer E.M Southern, a British biologist. Southern blotting involves separation of restricted DNA fragments by electrophoresis and then transferred to a nitrocellulose or a nylon membrane, followed by detection of the fragment using probe hybridization.

Separated by electrophoresis is transferred from gel to a membrane which in turn is used as a substrate for hybridization analysis employing labeled DNA or RNA probes specific to target fragments in the blotted DNA. Southern hybridization helps to detect specific fragment against a background of many other restriction fragments. Southern blotting is a technique which is used to confirm the identity of a cloned fragment or for recognition of a sub-fragment of interest from within the cloned DNA, or a genomic DNA. Southern blotting is a prerequisite to techniques such as restriction fragment length polymorphism (RFLP) analysis.

3-4.2.1 Procedure:

1. The high-molecular-weight DNA strands are fractioned using restriction enzymes.

2. The DNA fragments are separated based on size by agarose gel electrophoresis.

3. The gel with the restricted fragments is then laid on a filter paper wick which serves as a connection between the membrane and the high salt buffer.

4. The nitrocellulose membrane is placed on top of the gel and a tower of filter papers is used to cover it and these are kept in place with a weight. The capillary action drives the buffer soaking through the filter paper wick, through the gel and the membrane and into the paper towels. Along with the buffer passing through the gel the DNA fragments are also carried with it into the membrane and they bind to the membrane. This causes an effective transfer of fragments (up to 15 kb in length taking around 18hours or overnight.

For DNA fragments larger than 15 kb, before blotting an acid such as diluted HCl is used to treat the gel that depurinates the DNA fragments causing breakage of DNA into smaller pieces, resulting in more efficient transfer from the gel to membrane.

Now a days blotting is also done by applying electric field. This **electro blotting** technique depends upon current and transfer buffer solution to nucleic acids onto a membrane. Following electrophoresis, a standard tank or semi-dry blotting transfer system is set up. A stack is put together in the following order from cathode to anode: sponge, three sheets of filter paper soaked in transfer buffer gel, PVDF or nitrocellulose membrane, three sheets of filter paper soaked in transfer buffer and then again sponge.

Importantly the membrane should be located between the gel and the positively-charged anode, as the current and sample will be moving in that direction. Once the stack is prepared, it is placed in the transfer system, and suitable current is applied for a specific period of time according to the materials being used.

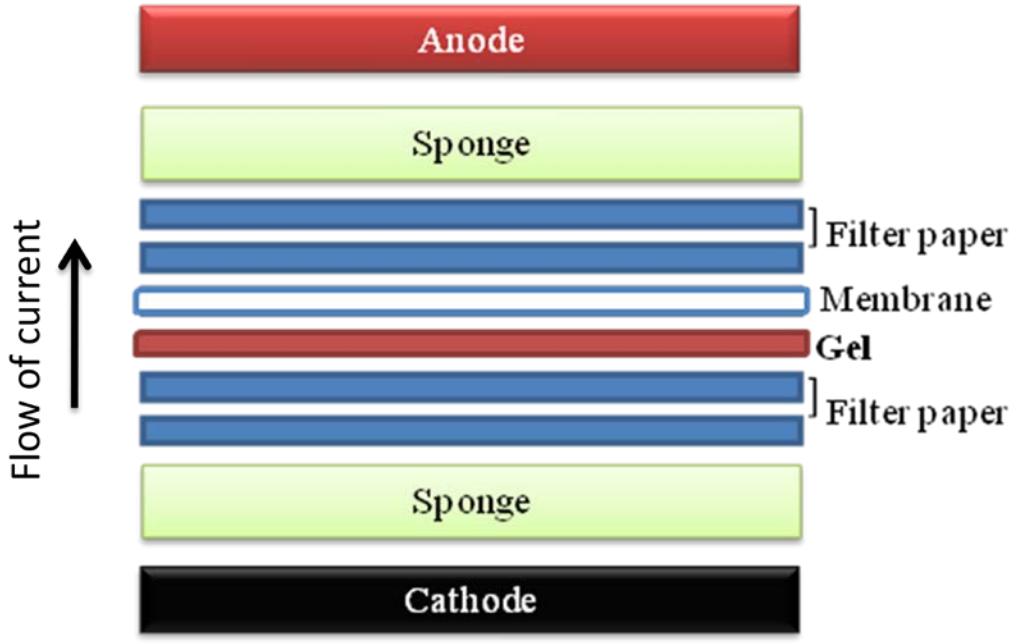


Fig 3-4.2.1: Set up for Electro-blotting system

5. For using alkaline transfer methods, the DNA gel is placed into an alkaline solution (like that of sodium hydroxide) causing denaturation of the double-stranded DNA. Denaturation in an alkaline environment enhances the binding between the negatively charged DNA and the positively charged membrane, causing separation to single DNA strands for further hybridization to the probe, alongside destroying any residual RNA that may persist in DNA. The membrane is washed with buffer to remove unbound DNA fragments.

6. The membrane which contains the transferred fragments is heated in presence or absence of vacuum at 80 °C for 2 hours or exposed to ultraviolet radiation (nylon membrane) for permanent attachment of the transferred DNA to the membrane.

7. The obtained membrane is then hybridized with a probe (a DNA fragment with a specific sequence whose presence in the target DNA is to be determined).

8. Labeling of the probe DNA is done for easy detection, usually radioactivity is incorporated or the molecule is tagged with a fluorescent or chromogenic dye. The hybridization probe may be made of RNA, instead of DNA in some cases where the target is RNA specific.

9. Washing of the excess probe from the membrane is done by using saline sodium citrate(SSC buffer) after the hybridization step and the hybridization pattern is studied on an X-ray film by autoradiography (for a radioactive or fluorescent probe), or via color development on membrane if a chromogenic detection method is employed.

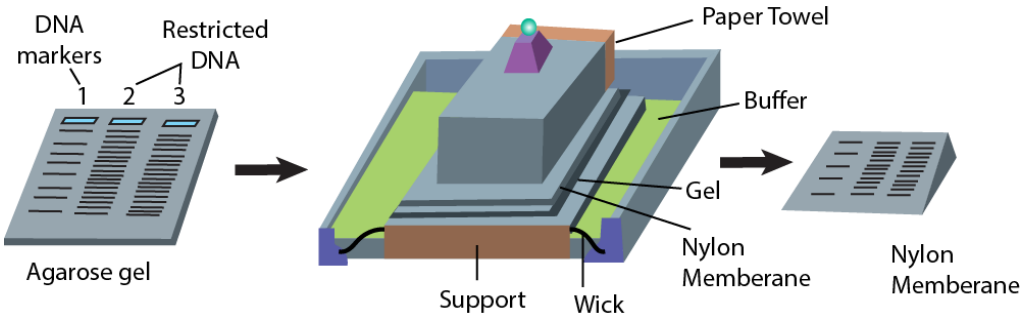


Fig 3-4.2.1: Steps of Southern Hybridization

3-4.2.2 Analysis of Southern Blot:

Hybridization of the probe to a specific DNA fragment on the membrane indicates the presence of a complementary fragment in the DNA sequence. Southern hybridization performed by digestion of genomic DNA using a restriction enzyme digestion, helps in determining the number of sequences (or gene copies) in the genome. For a probe hybridizing to a single DNA segment that has not been cut by the restriction enzyme, a single band is observed and on the other hand multiple bands will likely be observed when the hybridization occurs between the probe and several highly similar target sequence (Due to sequence duplication). Alterations in the hybridization conditions like enhancing the hybridization temperature or decreasing salt concentration, helps in altering specificity and hybridization of the probe to sequences that are less than 100% similar.

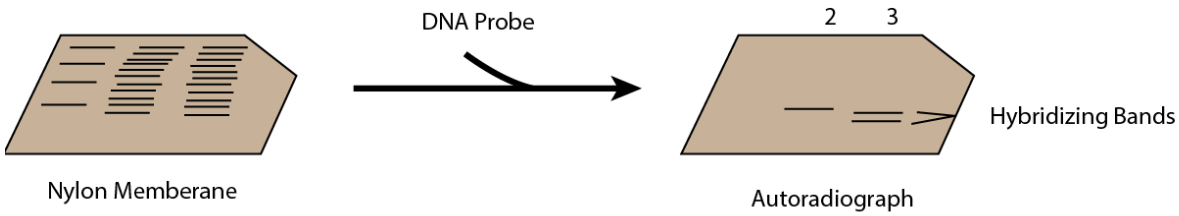


Fig 3-4.2 .2: Southern hybridization analysis

3-4.2 .3 Applications:

Southern blotting has been exploited for various applications which include:

a) Clone identification: One of the most common applications of Southern blotting is identification and cloning of a specific gene of interest. Southern blotting is carried out for identification of one or more restriction fragments that contain the gene of interest in genomic DNA.. After cloning and tentative identification of the desired recombinant by employing colony or plaque hybridization, southern blotting is further is used to confirm the clone identification and possibly to locate a shorter restriction fragment, containing the sequence of interest.

b) Restriction fragment length polymorphism Analysis: Another major application of Southern hybridization is restriction fragment length polymorphism (RFLP) mapping, which is crucial in construction of genome maps.

3-4.3 Northern hybridization:

Northern blotting was developed by James Alwine, George Stark and David Kemp (1977). Northern blotting drives its name because of its similarity to the first blotting technique, which is Southern blotting, named after the biologist Edwin Southern. The major difference is that RNA being analyzed rather than DNA in the northern blot.

Expression of a particular gene can be detected by estimating the corresponding mRNA by Northern blotting. Northern blotting is a technique where RNA fragments are separated by electrophoresis and immobilized on a paper sheet.Identification of a specific RNA is then done by hybridization using a labeled nucleic acid probe. It helps to study gene expression by detection of RNA (or isolated mRNA) in a sample.

In Northern blotting, probes formed of nucleic acids with a sequence which is complementary to the sequence or to a part of the RNA of interest. The probe can be DNA, RNA or chemically synthesized oligonucleotides of minimum 25 complementary bases to the target sequence.

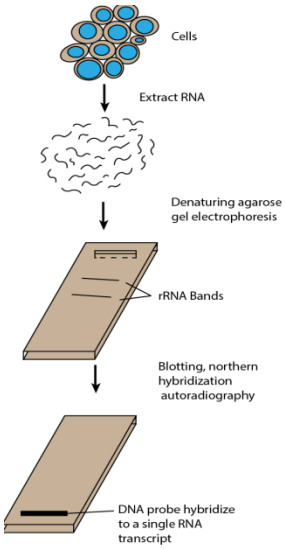


Fig 3-4.3:Steps of Northern Hybridization

3-4.3.1 Procedure:

The northern blotting involves the following steps:

1. Total RNA is extracted from a homogenized tissue sample or cells. Further eukaryotic mRNA can then be isolated by using of oligo (dT) cellulose chromatography to isolate only those RNAs by making use of a poly A tail.
2. The isolated RNA is then separated by gel electrophoresis.
3. The RNA samples separated on the basis of size are transferred to a nylon membrane employing a capillary or vacuum based system for blotting.

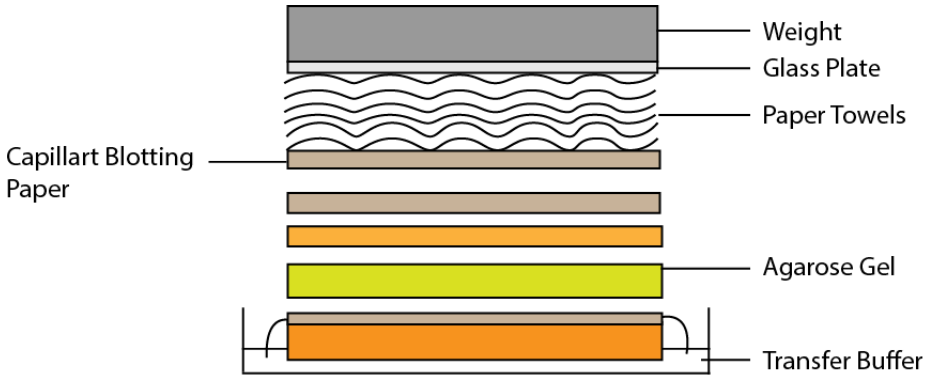


Fig 3-4.3.1: Setup for Northern blotting

4. Similar to Southern blotting, the membrane filter is revealed to a labeled DNA probe that is complementary to the gene of interest and binds.
5. The labeled filter is then subjected to autoradiography for detection.

The net amount of a specific RNA in a sample can be estimated by using Northern blot. This technique is widely used for comparing the amounts of a particular mRNA in cells under different conditions. The separation of RNA samples is often done on agarose gels containing formaldehyde as a denaturing agent as it limits the RNA to form secondary structure.

3-4.3.2 Analysis of Northern Blot:

RNA extract is electrophoresed in an agarose gel, using a denaturing electrophoresis buffer (containing formaldehyde) to ensure that the RNAs do not form inter- or intra-molecular base pairs, as base pairing would affect the rate at which the molecules migrate through the gel. After electrophoresis, the gel is blotted onto a nylon or nitrocellulose membrane, and hybridized with a labeled probe. If the probe is a cloned gene, the band that appears in the autoradiograph is the transcript of that gene. The size of the transcript can be determined from its position within the gel, and if RNA from different tissues is run in different lanes of the gel, then the possibility of differentially expressed gene can be examined.

3-4.3.3 Applications:

Northern blotting helps in studying gene expression pattern of various tissues, organs, developmental stages, pathogen infection, and also over the course of treatment. It has been employed to study overexpression of oncogenes and down-regulation of tumor-suppressor genes in cancerous cells on comparison with healthy tissue, and also for gene expression of immune-rejection of transplanted organ.

The examination of the patterns of gene expressions obtained under given conditions can help determine the function of that gene.

Northern blotting is also used for the analysis of alternate spliced products of same gene or repetitive sequence motif by investigating the various sized RNA of the gene. This is done when only probe type with variation in one location is used to bind to the target RNA molecule.

Variations in size of a gene product may also help to identify deletions or errors in transcript processing, by altering the probe target that can be used along the known sequence and make it possible to determine the missing region of the RNA.

3-4.4 Colony Hybridization:

It is a rapid method of isolating a colony containing a plasmid harboring a particular sequence or a gene from a mixed population. The colonies to be screened are first replica-plated on to a nitrocellulose filter disc that has been placed on the surface of an agar plate prior to inoculation. Master plate is retained for reference set of colonies. The filter bearing the colonies is removed and treated with alkali so that the bacterial colonies are lysed and the DNA they contain is denatured. The filter is then treated with proteinase K to remove protein and leave denatured DNA bound to the nitrocellulose. The DNA is fixed firmly by baking the filter at 80°C. A labeled probe is hybridized to this DNA which is monitored by autoradiography. A colony whose DNA print gives a positive auto radiographic result on X-ray film can then be picked from the reference plate.

Colony hybridization can be used to screen plasmid or cosmid based libraries.

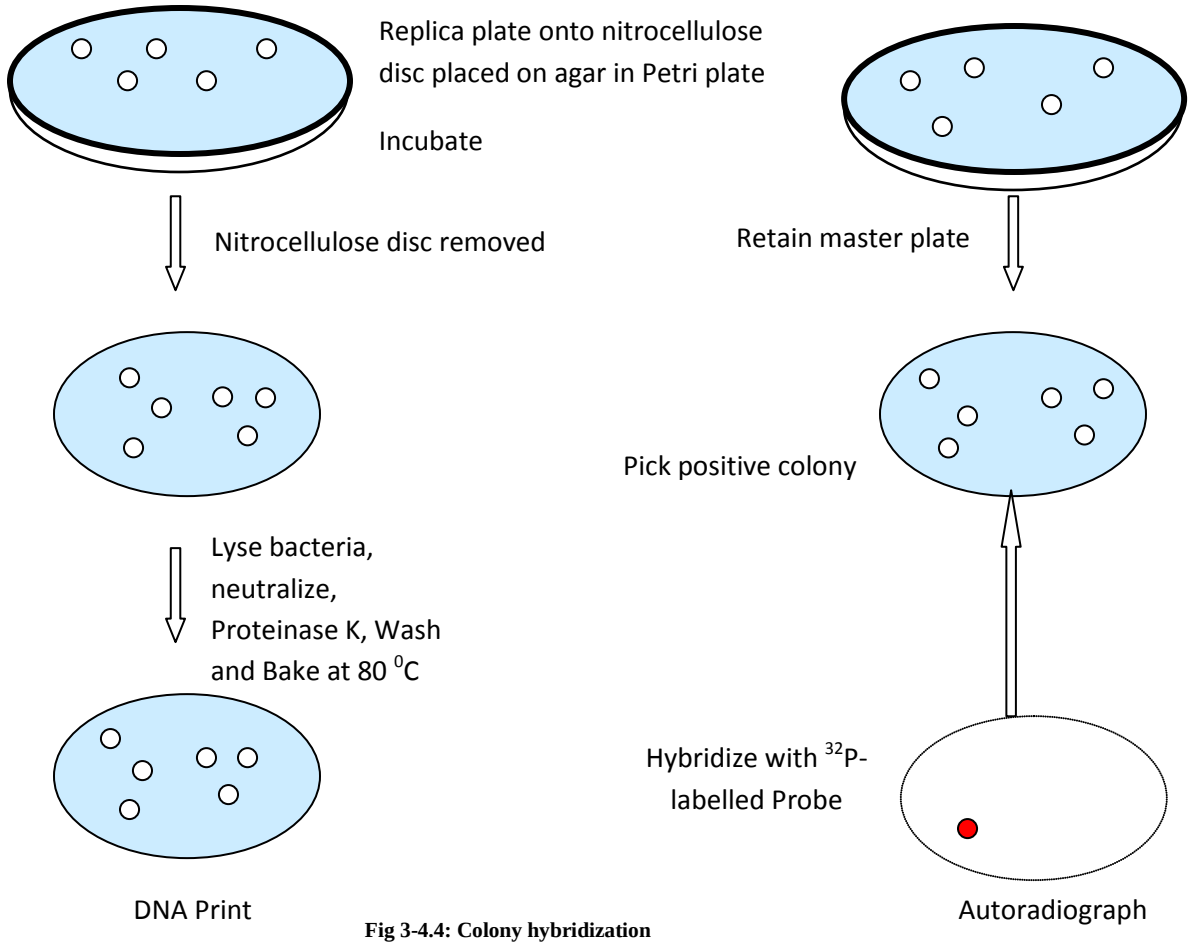


Fig 3-4.4: Colony hybridization

3-4.5 *In Situ* Hybridization (ISH)

It is a technique that employs a labeled complementary nucleotide strand (i.e. probe) for localizing specific DNA or RNA sequence targets within fixed tissues and cells (i.e *in situ*). Probes used for hybridization can be double-stranded DNA probes, single-stranded DNA probes, RNA probes, synthetic oligonucleotides. There are two ways available to detect DNA or RNA targets

Chromogenic (**CISH**) in situ hybridization and Fluorescence in situ hybridization (**FISH**).

3-4.5.1 Chromogenic in situ hybridization (CISH)

It uses the labeling reactions involving alkaline phosphatase or peroxidase reactions to visualize the sample using bright-field microscopy. It is primarily used in molecular pathology diagnostics. CISH can also be employed for samples like fixed cells or tissues, blood or bone marrow smears and metaphase chromosome spreads.

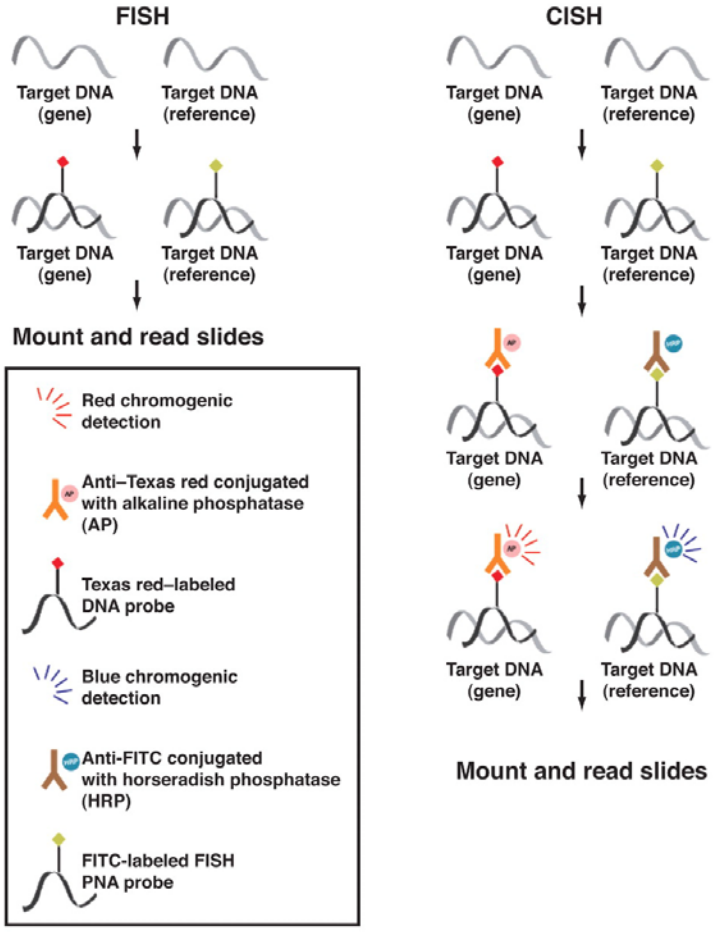


Fig.3-4.5.1 Use of dual-color Chromogenic in situ hybridization (CISH) in combination with fluorescence in situ hybridization (FISH) probes. FITC, fluorescein isothiocyanate; PNA, peptide nucleic acid.

(Obtained from http://ajcp.ascpjournals.org/content/133/2/205/F1.expansion.html)

3-4.5.2 Fluorescence in situ hybridization (FISH)

FISH is a cytogenetic technique that uses fluorescent probes that bind to complementary targets and sample is visualized using epi-fluorescence or confocal microscopy. Using differently labeled probes, we can visualize several targets in a single sample. It is used for spatial-temporal patterns of gene expression and resolving genetic elements in chromosomal preparations.

Cells or tissues to be analyzed are fixed and permeabilized with Proteinase K to allow target accessibility. Probe is constructed and tagged using non-radioactive labels like biotin, digoxigenin or fluorescent dye (FISH). Probe must be large enough to hybridize specifically with its target. Probe is applied to fixed sample and incubated to several hours to allow hybridization. Washing is done to remove non-specific or unbound probe removal. Results are then visualized using either bright-field or confocal microscopy.

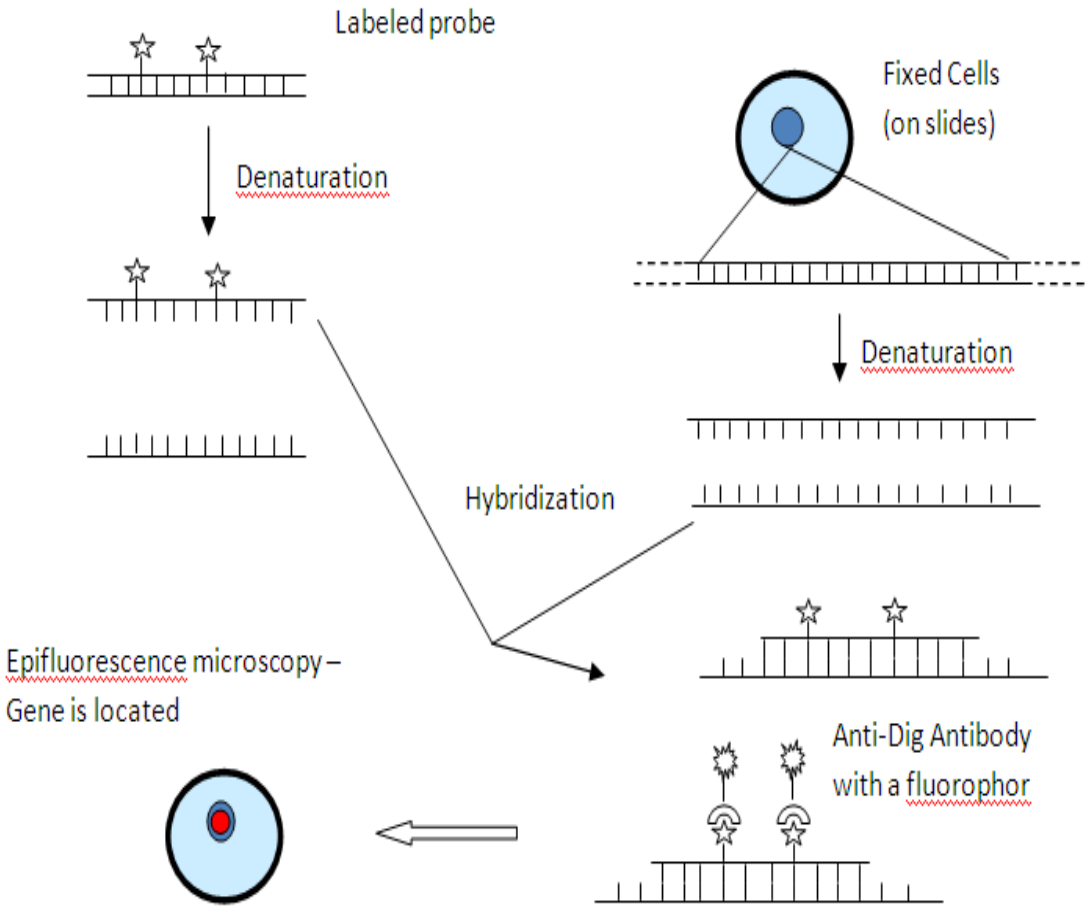


Fig3-4.5.2: Fluorescent in-situ hybridization

3-4.6 Dot Blot and Slot Blot Hybridization

These two techniques represents the simplification of Southern and Western blots saving the time involved in procedures of chromatography, electrophoresis, restriction digestion and blotting of DNA or proteins from the gel to membrane. Here nucleic acid mixture is directly applied (blotted) on to the nylon or nitrocellulose membrane where hybridization between probe and target takes place, denatured to single-stranded form and baked at 80°C to bind DNA target to membrane. In dot-blot, target is blotted as circular blots whereas in slot-blots, it is in the form of rectangular blots. Due to this, slot-blot offers greater precision in observing different hybridization signals. After blotting, membrane is allowed to dry and non-specific sites are blocked by soaking in blocking buffer containing BSA. It is then followed by hybridization of labeled probe for detection of specific sequences or gene.

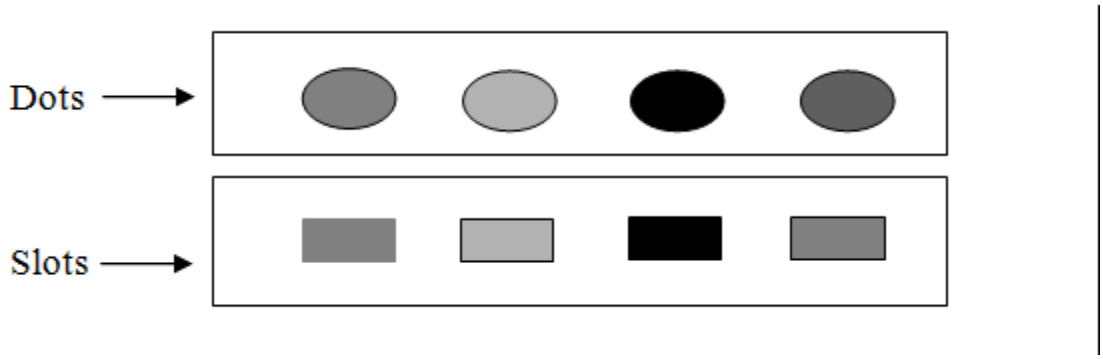


Fig 3-4.6 Dots and slots in dot and slot blot hybridization

These procedures can only detect presence and absence of particular sequence or gene. It cannot distinguish between two molecules of different sizes as they appear as single dot on membrane. It also has application in detecting alleles that differ in single nucleotide with the help of allele-specific oligonucleotides.

Bibliography

H. Lodish, A.Berk, P. Matsidaira, Chris A. Kaiser(2004); Modern Cell Biology, 5th edition; Macmillan Higher Education.

Sandy B. Primrose & Robert W. Old (2001); Principle of Gene Manipulation, 6th edition; *Blackwell Scientific Publ.*

T.A.Brown (2002); Genomes, second edition; *Oxford: Wiley-Liss press*.

MODULE 3: LECTURE 5

PROBE AND TARGET SEQUENCES

3-5.1 Introduction

Probes are short section of DNA or RNA with an additional tagged or labeled chemical entity that are used to bind to its complimentary strand and thereby allows detection of candidate nucleic acid molecule. This chemically synthesized entity can be a fluorescent molecule or it can be an attachment to a colored bead, or quantum dots (Cd-Se Qdots, Zn-SQdots), photochromic compounds, isotopic labeling, non-isotopic labeling etc. It allows us to visualize when a probe attaches to DNA, RNA or other target nucleic acids.

3-5.2 Homoduplex and Heteroduplex between probe and target:

For hybridization, the interacting single-stranded nucleic acid molecules should have a sufficiently high degree of base complementarity. It involves interaction of single strands of two sources of nucleic acids:

1. **Probe** consists of either chemically synthesizednucleic acids or modified oligonucleotides with known sequence that helps to identify similar or identical complementary sequences.

2. **Target** consists of template nucleic acid molecules with which probes will hybridize and form complex and heterogeneous mixture.

Probe or target can be in double-stranded form which must be separated before hybridization by heating or by alkaline treatment. Complementary base pairs are allowed to re-associate when single stranded probe is mixed with single stranded target. Various types of re-association can occur at this stage. Complementary probes and complementary targets can hybridize to each other forming homoduplexes. Both Homo and hetero-duplexes are useful for hybridization assay to detect a particular sequence in target.

The whole methodology revolves around the principle of hybridizing probe with a known sequence to a target which may not have related sequences to anneal with. Conditions of hybridization must be stringent enough to eliminate mismatched heteroduplexes. Low concentration of salts and high temperature increases the stringent conditions. If the probe used is small enough then hybridization reactions can be chosen such that heteroduplex is unstable even when there is single mismatch.

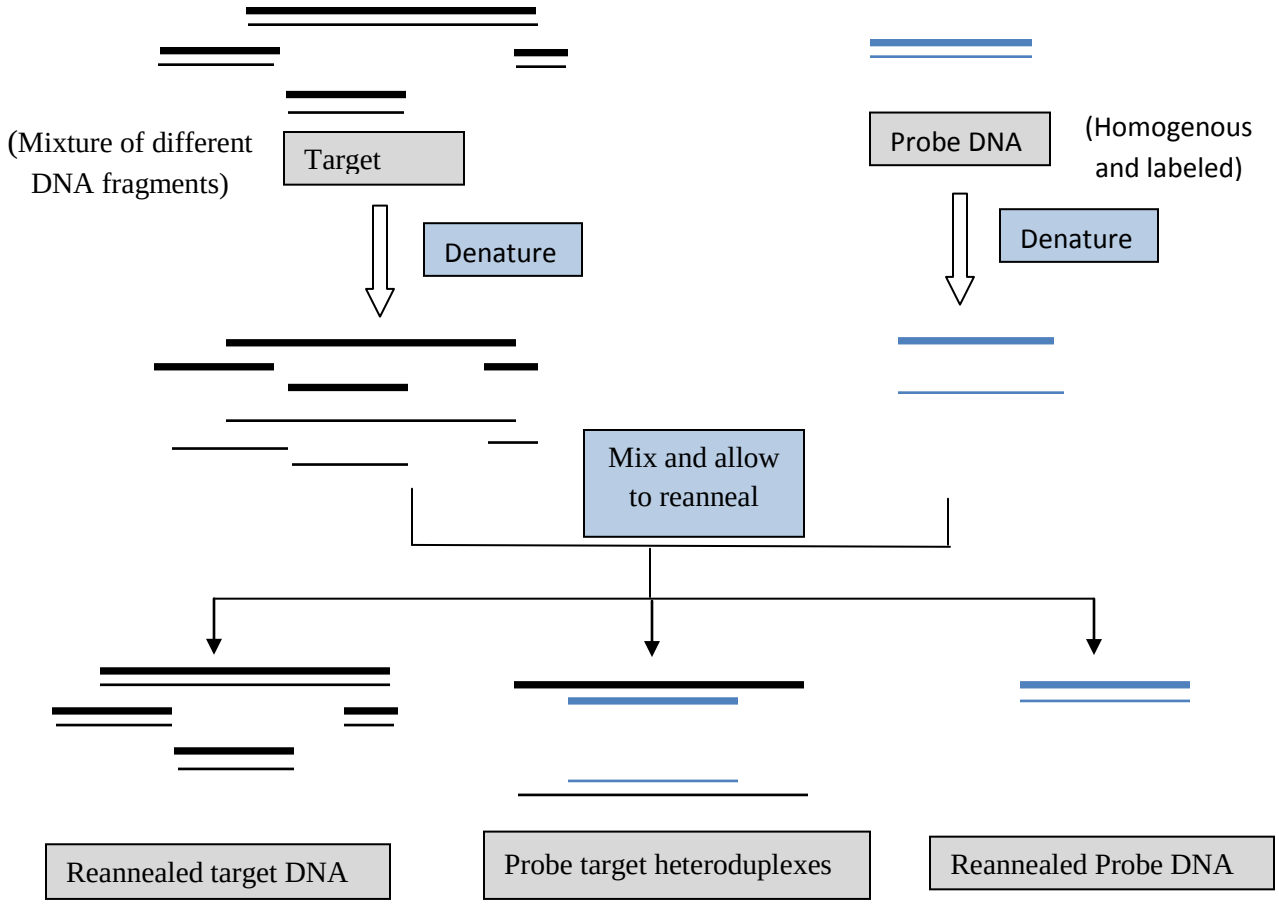


Fig 3-5.2: Method of Nucleic acid hybridization

Probe and target are first denatured and annealed forming the probe-probe homoduplexes and target-target homoduplexes, but the desired reaction is probe-target heteroduplex formation.

3-5.3 Gene Probes

Gene probes are generally longer than 500 bases and consist all or most of a target gene. They can be generated in two ways. Cloned probes are normally used when a specific clone is available or when the DNA sequence is unknown and to be cloned first in order to be mapped

and sequenced. Polymerase chain reaction is a powerful tool for making gene probes because it is possible to amplify and label simultaneously. Having the whole sequence of a gene, which can easily be obtained from Databases (Genbank, EMBL, DDBJ) primers can be designed to amplify the whole gene orgene fragments. It is more convenient when the gene of interest is PCR amplified because there is no need for restriction enzyme digestion, electrophoresis and elution of DNA fragments from vectors. If PCR amplification gives nonspecific bands, it is recommended to gel purify the specific band that will be used as a probe.

Gene probes generally provide greater specificity than oligonucleotides because of their longer sequence and because more detectable groups per probe molecule can be incorporated

into them than into oligonucleotide probes.

3-5.4 Hybridization Probes

Probe is labeled in *standard hybridization assay* whereas target is labeled in *reverse hybridization assays*.Ease of preparation and reliability of labeled probes make the *standard hybridization* assays as preferable choice over reverse hybridization methods. Nucleic acid probes can be single-stranded or double-stranded molecules, but the working probe must be single-stranded. The probes,whichmay constitute either short sequences of DNA, RNA or synthetic oligonucleotides, are used in hybridization based techniques such as Southern, Northern blotting, colony and plaque hybridization, in situ hybridization and sequencing by hybridization.

Origin and characteristics of nucleic acid hybridization probes:

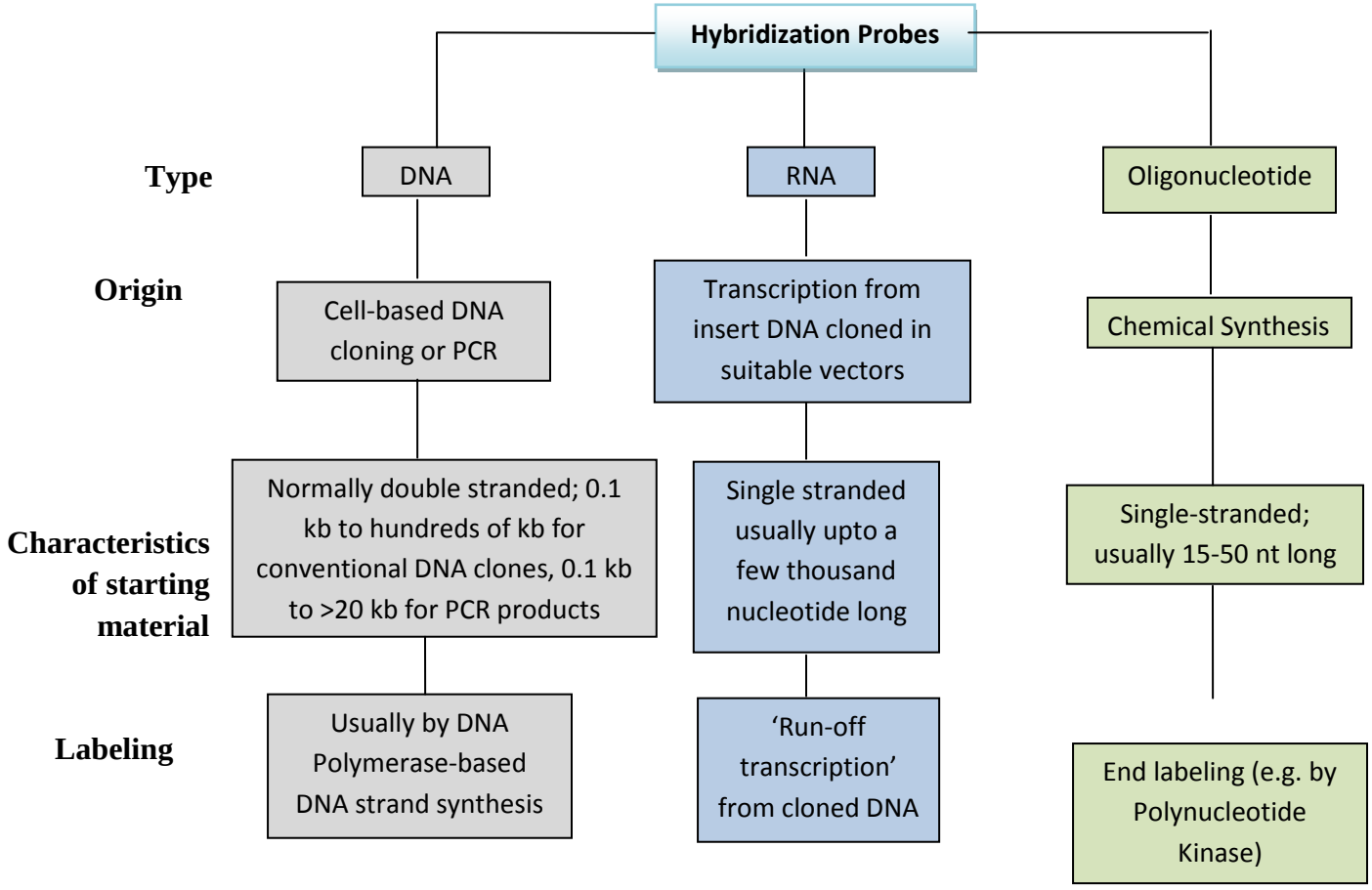


Fig. 3-5.4 Features of Hybridization Probes

3-5.4.1 DNA probes:

Conventional DNA probes are isolated by cell-based DNA cloning or by PCR. In cell-based DNA cloning, the starting DNA can range from 0.1 kb to a range of ~10kb in length and is often double-stranded. PCR-derived DNA probes are usually 10 kb long and are usually double-stranded. These DNA probes are labeled using labeled dNTPs during in vitro synthesis.

3-5.4.2 RNA probes:

Detection of specific nucleic acid sequences can be accomplished by hybridization with a labeled RNA probe. RNA probes are sequences of a variable length that are used to detect the presence of complementary nucleotide sequences in a sample. RNA probes are first labeled with radioactive nucleotides or with modified nucleotides that can be detected by fluorescence or chemiluminescence. They can be used for Northern blotting, RNase protection assays, Southern blotting, downstream of polymerase chain reaction (PCR), and *in situ* hybridization analysis.

3-5.4.2 Synthetic probes:

These probes are chemically synthesized which can generate short single stranded stretches 15–50 bases in length. They are usually highly specific as they are designed particularly for target DNA. Probes with degenerate sequences can also be synthesized using protein sequences. Degeneracy is introduced when parallel synthesis of oligonucleotides is done with similarity at some sites and difference at others. These probes are usually labeled using ^{32}P isotope or other labeled group at the 5′ end.

3-5.5 Labeling Of Probes:

Probes can be labeled at specific location within the oligonucleotides or internally at multiple sites. Some probes are of defined length and some are heterogeneous populations of labeled molecules.

In vivo labeling: DNA and RNA can be directly labeled inside tissue culture cells by adding labeled deoxynucleotides in culture plate in vivo. This method is restricted only to prepare labeled viral DNA from virus-infected cells and to study RNA processing events.

In vitro labeling: It is a more versatile method involving *in vitro* labeling of purified RNA, DNA or oligonucleotide using DNA polymerasesfor incorporation of labeled nucleotides.

3-5.5.1 Labeling during synthesis:

*In vitro*labeling of DNA can be done by various methods as follows:

a) Nick-translation

b) Random primed labeling

c) PCR-mediated labeling-Labeling of RNA is generally accomplished by an*in vitro* transcription system.These procedures require DNA or RNA polymerases to add labeled nucleotides to synthesize in vitro probes. It requires at least one labeled nucleotides among four nucleotides.

3-5.5.1.1 DNANick translation:

It involves insertion of random single-strand breaks called nicks in one of the strands of double stranded target DNA which exposes 3′- OH termini and 5′-PO_4termini.The nicks are introduced by endonuclease like pancreatic deoxyribonuclease I (DNase I).

Fig. 3-5.5.1.1 DNA Nick translation by DNase I

Addition of DNase I and multi-subunit enzyme *E. coli* DNA polymerase I is used for nick translation which contributes both activities like:

(i) 5′ → 3′ exonuclease that attacks the exposed 5′ termini of a nick and sequentially removes the nucleotides in 5′ → 3′ direction;

(ii) DNA polymerase adds nucleotides to the free 3′-OH group, in 5′ → 3′ direction replacing the nucleotides removed by exonuclease and causing lateral displacement (translation) of the nick.

This method requires 100-fold less radioactive precursor than in-vivo labeling method. The amount of radiolabel incorporated depends on number of nicks created by DNase I. Too much or too little nicking must be avoided. At low temperatures (about 15° C), the

disadvantage is that only one complete regeneration of existing nucleotide sequence takes place and reaction does not proceed further.

3-5.5.1.2 Random primed DNA labeling:

This method known as oligo-labeling is based upon hybridization of a mixture of all possible hexanucleotides. The template DNA is initially denatured and then cooled slowly so as to allow random hexanucleotides to bind at complementary sequences at which extension takes place through PCR.

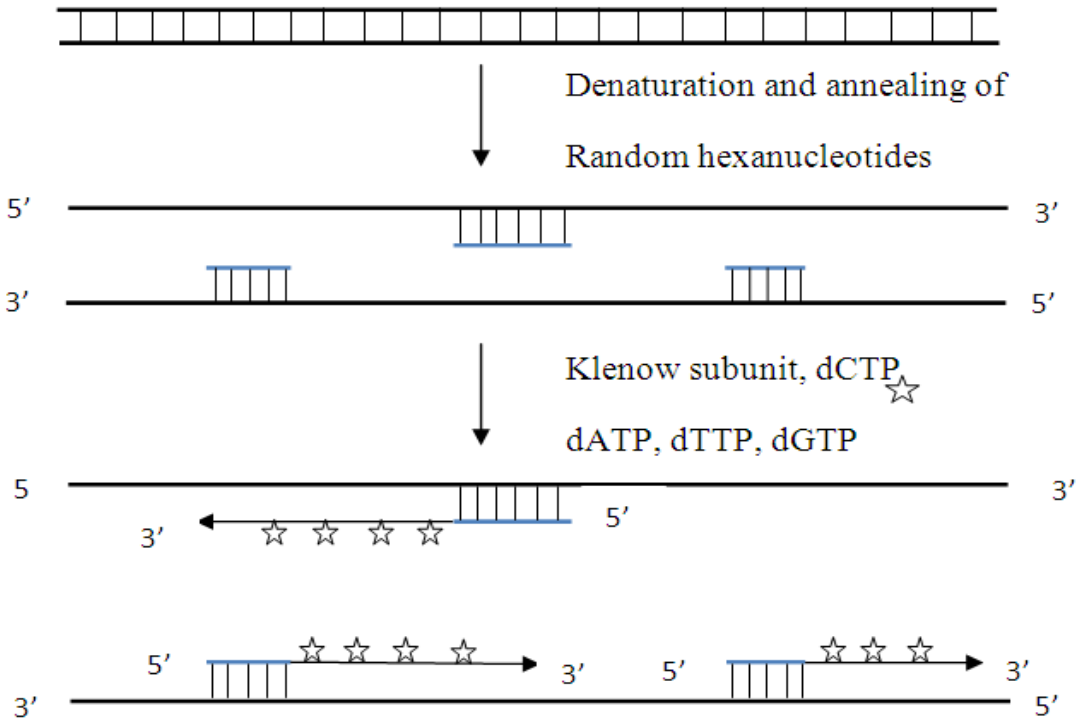


Fig3-5.5.1.2 Random primed DNA labeling

The synthesis of new complementary DNA strands is primed by bound hexanucleotides and catalyzed by Klenow subunit of DNA polymerase I (contains the polymerase activity in the absence of associated 5'→3' exonuclease activities, sometimes 3'→5' exonuclease activity also removed for the reaction).This method produces labeled DNAs of high specific activity. Since binding of primer to template DNA is random and primer represents all possible sequence combinations,uniform labeling of DNA occurs.

The length of random primers is crucial. Primers shorter than 6 bases are very poor primers, whereas those longer than 7 have higher tendency to self–anneal and cause primer duplex. Thus, probe should be either 6 or 7 nucleotides in length.

Random priming is inherently simpler than nick translation because the requirements for two nuclease activities are eliminated. The probes generated by random priming method are more homogenous in size and behave more reproducibly in hybridization reactions. Average size of probe cannot be controlled with great accuracy in nick translation method.

Template DNA used preferably should not be closed, circular ds DNA because they are inefficient templates. Shorter templates generate probes of low specific activity that may not hybridize under stringent conditions.

3-5.5.1.3 PCR mediated DNA labeling:

This method has several advantages over other methods.

- Defined segments of target DNA can be amplified and labeled independently of restriction sites.
- Amount of template DNA is required very small and
- No need to isolate fragments of DNA or to sub-clone into vectors containing bacteriophage promoters.

The standard PCR reaction can be modified to incorporate labeled nucleotides. The methods commonly used are:

- **Standard PCR-based DNA labeling-** The probe generation reaction is modified to incorporate one or more labeled nucleotide precursors at a concentration same as oligonucleotide conc.(K_m) or slightly above K_m and others at concentrations exceeding K_m, which become incorporated into the PCR product throughout its length.
- **Primer-mediated 5′ end labeling**. This method uses a 5’ end-labeled primer, which is incorporated during PCR reaction. It is utilized in DNA sequencing, PCR based mutagenesisetc.

Radiolabeled probes can be generated for both strands using equal concentrations of primers or biased heavily in favor of one strand of DNA using higher concentration of one primer.

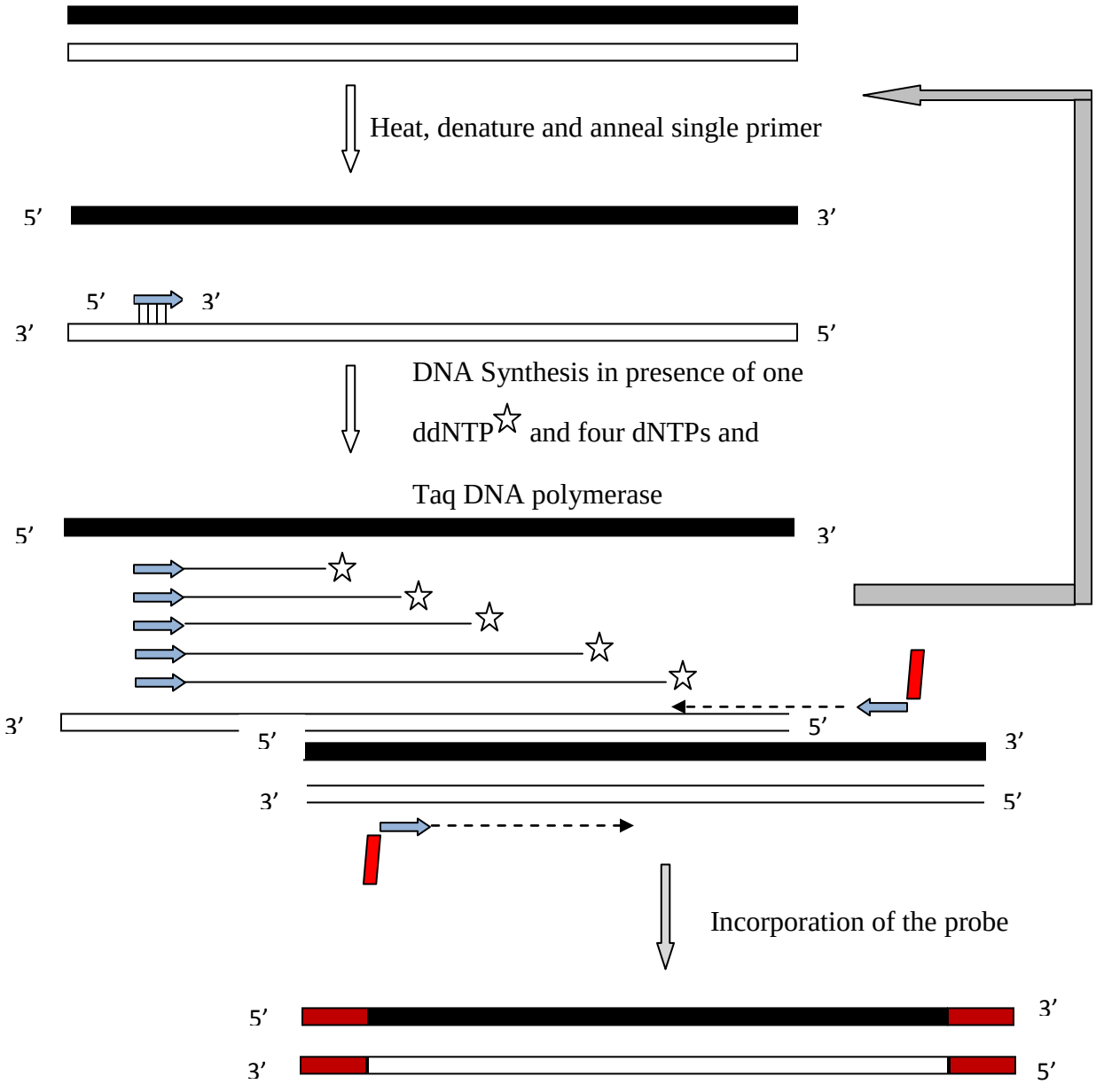
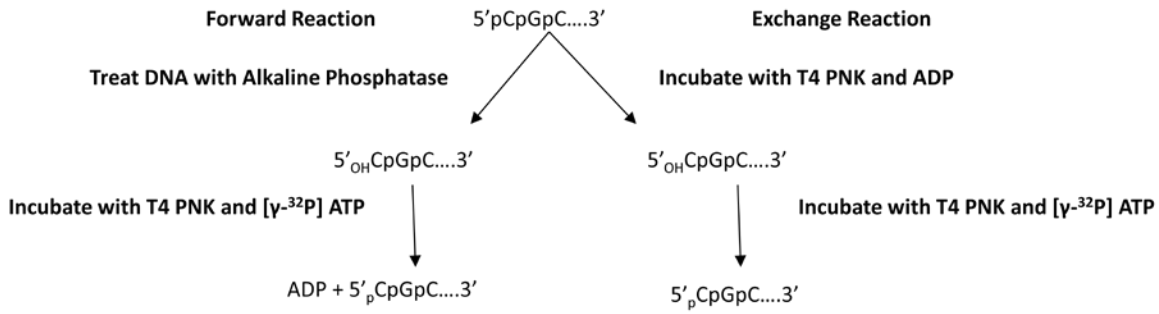


Fig3-5.5.1.3.Standard PCR-based DNA labeling

3-5.6 Labeling of ss DNA, ssRNA and ds DNA with protruding 5’ terminus or blunt ended or recessed 5’ termini:

Oligonucleotides are usually end-labeled using polynucleotide kinase. The label is in the form of a ^{32}P isotope at the γ-phosphate position of ATP. The polynucleotide kinase catalyzes an exchange reaction with the 5′-terminal phosphates at high concentration of dADP and low concentration of dATP. Alkaline phosphatase such as calf intestinal phosphatase (CIP) or shrimp alkaline phosphatase (SAP) can be used to remove 5’ phosphate group and polynucleotide kinase (PNK) is used to add labeled phosphate using labeled dATP.



The same methodology is used for labeling double stranded DNA. Fragments which carry label at one end can only be generated using restriction enzyme which cleaves at internal sites, thus producing two fragments which can be separated by gel electrophoresis and purified. Fill-in end labeling is a popular approach to label large DNA fragments using Klenow subunit of E. coli DNA polymerase.

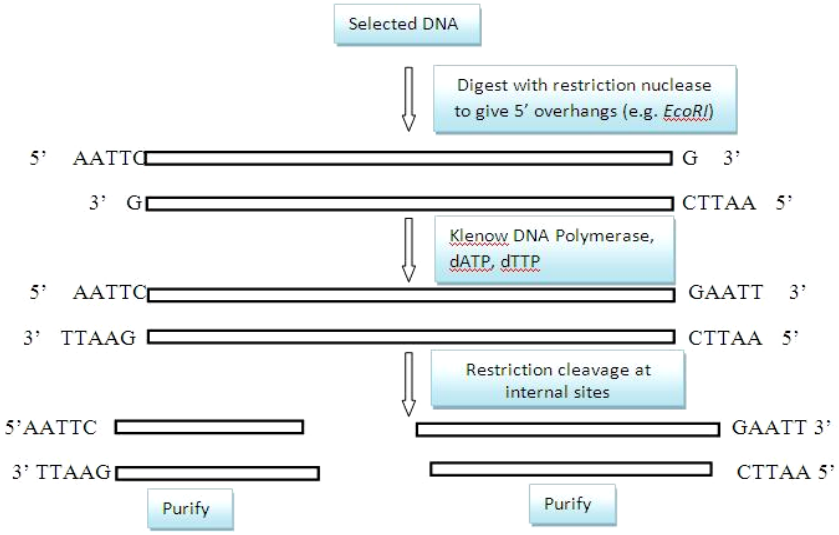


Fig3-5.6: Labeling of double-stranded DNA

3-5.6.1 Labeling 3' Termini of dsDNA with recessed 3' termini or blunt-ended:

Klenow fragment of *E.coli* DNA polymerase I is used to incorporate ^{32}P dNTPs into a recessed 3' terminus generated by restriction digestion. The choice of ^{32}P dNTP for labeling reaction depends on sequence of protruding 5' termini of DNA. End fragments created by restriction endonucleases can be labeled with ^{32}P dATP.

$$5'\text{-}G_{OH}\ 3' \quad + \quad 5'_{P}\ AATTC\text{-}3' \quad \longrightarrow \quad 5'\text{-}GAA_{OH}\text{-}3' \quad + \quad 5'_{P}AATTC\text{-}3'$$

$$3'\text{-}CTTAA_{p}5' \quad + \quad 3'_{OH}G\text{-}5' \quad \longrightarrow \quad 3'\text{-}CTTAA_{p}5' \quad + \quad 3'_{OH}\ AAG\text{-}5'$$

The terminal nucleotide of blunt end DNA fragment can be replaced by weak 3'→5' exonuclease and strong polymerase activity of klenow fragment.

3-5.6.2 Labeling 3'termini of dsDNA with protruding 3'termini:

Bacteriophage T4 DNA polymerase have strong 3'→5' exonuclease activity than of klenow fragment. It is used to digest 3'-protruding termini and then continues at slower pace to remove 3' nucleotides from ds DNA. At higher concentration of dNTPs, radiolabelled dNTP is added by polymerase activity.

Alternatively, protruding 3'termini can be labeled by calf thymus terminal transferase to catalyze the transfer of ^{32}P dideoxy ATP. Because it does not carry 3'-OH group, no additional nucleotides can be added to protruding 3'end.

3-5.6.3 Labeling RNA probes:

RNA probes yield stronger signals in hybridization reactions than DNA probes of equal specific activity which may be due to innately higher stability of hybrid involving RNA. They are the probes of choice when analyzing transcripts of mammalian genes.

RNA probes are most easily achieved by in vitro transcription in which insert DNA is cloned adjacent to multiple cloning sites in suitable plasmid expression vector. This expression vector has a phage promoter sequence which is recognized by corresponding phage RNA polymerase. In the given plasmid vector pSP64, bacteriophage SP6 promoter sequence is recognized by SP6 RNA polymerase which starts transcription from a particular initiation site after which insert has been cloned. Highly specific labeled RNA probes can be generated using radiolabeled dNTPs.

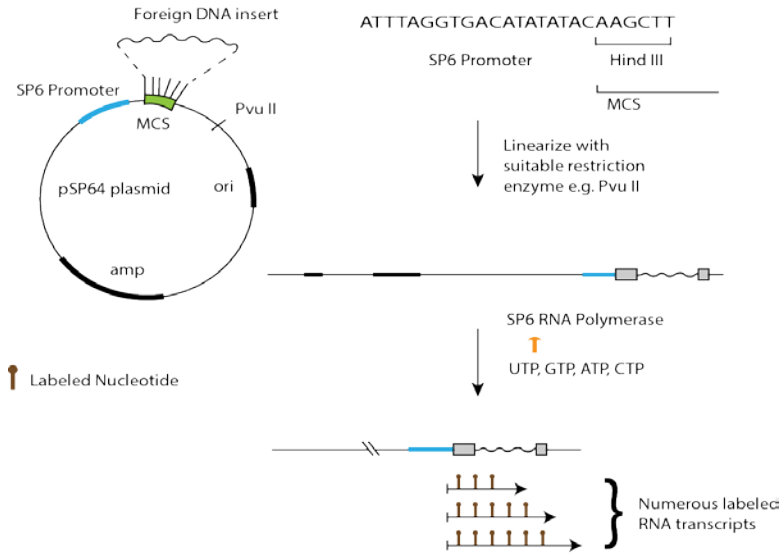


Fig3-5.6.3 Labeling RNA probes

Riboprobes are mostly created by run-off transcription in which RNA insert is cloned in specialized plasmid vectors. The 5' terminus of the transcript is fixed by bacteriophage promoter, but 3' terminus is defined by downstream site of cleavage by restriction enzyme. Restriction enzymes that generate blunt or 5'-protuding termini produce best linear templates. Enzymes generating protruding 3' termini, should be avoided because transcription of such templates results in synthesis of RNA molecules that are aberrantly initiated at the termini of templates.

Commonly used systems for generation of riboprobes are Bacteriophage T3 and T7 promoter/RNA polymerase systems. Labeled riboprobes, both sense and antisense can be generated from any gene cloned in such vectors (in either of the two orientations). These are widely used in tissue *in situ* hybridization.

The DNA fragment to be transcribed can also be amplified in PCR with primers having 5' ends encode synthetic promoters for bacteriophage RNA polymerases. Following purification, the PCR products can be used as double stranded templates for in-vitro transcription.

3-5.7 Preparation of Labeled Nucleotides:

Nucleotides can be labeled by **isotopic and non-isotopic** methods.

3-5.7.1 Isotopic labeling:

Isotopes generally used for labeling nucleotides are ^{32}P, ^{33}P, ^{35}S or ^{3}H. They can be detected directly in solution or on X-ray film using autoradiography.

Properties of radioisotopes used for labeling DNA and RNA probes:

Radioisotope	Half-life	Energy of emission
^{3}H	12.4 years	0.019 MeV
^{32}P	14.3 years	1.710 MeV
^{33}P	25.5 years	0.248 MeV
^{35}S	87.4 years	0.167 MeV

The strength of autoradiography signal depends on intensity of radiation emitted by radioisotope and duration of exposure. ^{32}P emits high energy β-particles which offer high detection sensitivity. Thus, it is widely used in Southern blot hybridization, dot-blot hybridization, colony hybridization. But it is relatively unstable and when fine resolution is required to interpret results, the image is unambiguous due to its high energy β-particle emission.Due to this,^{35}S-labeled, ^{33}P-labeled (moderate half-lives) and^{3}H-labeled nucleotides are used which emit less energetic β- radiation. They are used in DNA sequencing and in-situ hybridization.^{3}H requires long exposure time due to low energy β-particle emission.

3-5.7.2 Non-isotopic labeling:

Non-isotopic labeling systems involve the use of nonradioactive probes. These methods are developed recently as compared to radioisotope labeling methods, but are finding wide variety of applications in different ways. Two types of non-radioactive labeling are conducted: direct and indirect.

3-5.7.2.1 Directnon-isotopic labeling, where a nucleotide containing label such as Fluorescein, Texas Red, Rhodamine that will be detected when incorporated with the help of spacer molecule. These modified nucleotides having fluorophore tag, fluoresce when excited by light of certain wavelength.

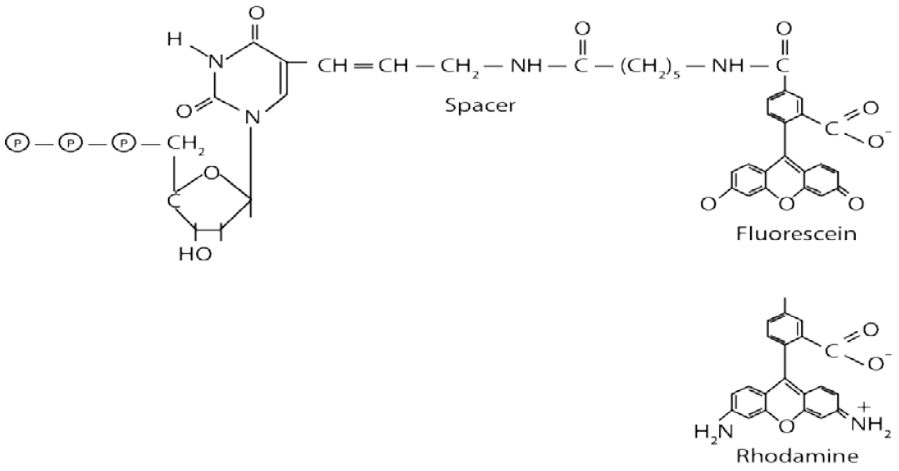


Fig 3-5.7.2.1 An example of fluorescein conjugated dUTP. The fluorescein group is linked to the 5′ carbon atom of the uridine by a spacer group. Similarly, Rhodamine can also be used in place of fluorescein.

3-5.7.2.2 Indirect non-isotopic labeling involves chemical linkage of reporter molecule to a nucleotide. When this modified nucleotide is incorporated into DNA, then it is specifically bound to a protein or other ligand which has high affinity against the reporter group. Long spacer is introduced between nucleotide and reporter so as to reduce steric hindrances for binding of affinity molecule.

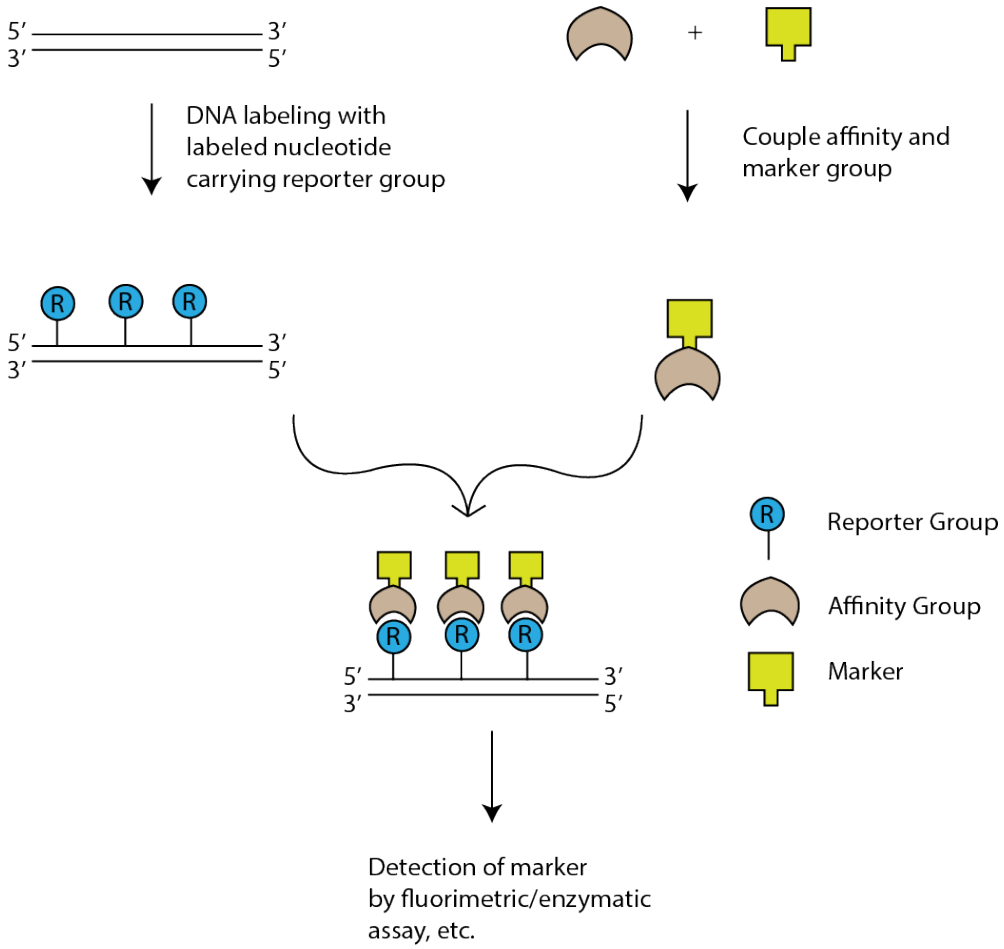


Fig. 3-5.7.2.2 Indirect non-isotopic labeling

Two widely used non-isotopic labeling methods are:

Biotin-Streptavidin Method: This method uses two ligands which has high affinity towards each other: Biotin works as the reporter and the bacterial protein streptavidin is used as the affinity molecule. Biotinylated nucleotides like bio-11-dUTP are used as labeling agents with a spacer of 4-16 C atoms long between biotin and dNTP. However, Biotin is a ubiquitous constituent of mammalian tissues and tends to stick easily to certain type of nylon membranes which leads to high levels of background during in

situ, northern and southern hybridization. To overcome this background problem, digoxigenin is used.

Digoxigenin, a plant steroid obtained from *Digitalis* plant and is used as a reporterand an affinity molecule.Digoxigenin is thus an all-purpose immuno-tag, and in particular a standard immunohistochemical marker for in situ hybridization.

Enzymatic methods are mostly used to label DNA probes with biotinylated nucleotides. Photochemical labeling of biotin to nucleic acids can also be used. The label is linked to nitrophenylazido group that is converted by UV irradiation to highly reactive nitrene that form stable covalent linkages to DNA or RNA.

End labeled DNA can be used as:

1. Molecular-weight standards in Southern blotting
2. Probes in gel-retardation experiments
3. Tracers for small quantities of DNAs on gels
4. Probes for screening bacterial colonies or plaques
5. Substrates for Maxam-Gilbert sequencing
6. Probes for RNA mapping with S1 Nuclease or Mung bean nuclease
7. Primers in primer-extension reactions.

3-5.8Detection of non-radioactively labeled probes after hybridization:

Affinity molecules (streptavidin or digoxigenin-specific antibody) are conjugated with a variety of marker groups or molecules. They include various fluorophores or enzymes such as alkaline phosphatase and peroxidase which can permit detection via colorimetric assays, chemical luminescence assays or fluorescent assay.

In **colorimetric assays**, alkaline phosphatase catalyzes removal of phosphate group from BCIP (5-bromo-4-chloro-3-indolyl phosphate), generating a product that dimerizes to di-bromo-di-chloro indigo, which reduces NBT (Nitrobluetetrazolium) to insoluble purple dye, diformazan that becomes visible at sites where probe has hybridized.

Fluorescent assays make use of HNPP (2-hydroxy-3-naphthoic acid 2'-phenylanilide phosphate). After de-phosphorylation by alkaline phosphatase, HNPP generates fluorescent precipitate on membranes that can be excited by irradiation at 290nm. The response signal emitted at 509nm are captured by CCD cameras.

Chemiluminescence is the fastest and most sensitive assay using HRP (Horseradish peroxidase) – luminal detection system. HRP catalyzes the oxidation of luminal in the presence of H_2O_2, generating reactive peroxide that emits light at 425nm during decomposition to its ground state.

Applications: This technique is used in physical mapping, karyotyping and phylogenetic analysis, gene expression profiling in developmental biology, pathogen profiling, and abnormal gene expression in pathology and for morphology and population structure of microorganisms.

BOOK

Strachan T and Andrew P Read. Human Molecular Genetics,2nd edition.

Samrook J and David W. Russell Molecular Cloning:A Laboratory Manual, 3rd Edition.

BIBLIOGRAPHY

Eleftherios P. Diamand, Theodore K. Christopoulos (1991). The Biotin-(Strept)Avidin System: Principles and Applications in Biotechnology. *CLIN. CHEM.* 37/5, 625-636.

Grunstein M., David S.,Hogness (1975). Colony hybridization: A method for the isolation of clofied DNAs that contain a specific gene.*Proc. Nat. Acad. Sci.* USA Vol. 72, No. 10, pp. 3961-3965.

BlouinJ. L., Rahmani Z., ChettouhZ., PrieurtM., FermaniantJ et al1990. Slot Blot Method for the Quantification of DNA Sequences and Mapping of Chromosome Rearrangements: Application to Chromosome 21.*Am. J. Hum. Genet.* 46:518-526.

Pennina R. Langer-Safer, Michael Levine, David C. Ward(1982). Immunological method for mapping genes on Drosophila polytene chromosomes. *Proc. NatL Acad. Sci.* USA Vol. 79, pp. 4381-4385.

Marilena Aquino de Muro. Probe Design, Production, and Applications. Medical Biomethods Handbook 2005, *pp 13-23.*

MODULE 3- LECTURE 6

NUCLEIC ACID MUTAGENESIS: *IN VIVO* AND *IN VITRO*

3-6.1 Introduction

Mutation of the genes in higher organisms can be carried out by site- specific or random variation in the gene expression cassette using the model microbes (eg. *E.coli*, *S.cerevisiae* etc). The modification of the sequence of the gene is done by introducing an alteration in the sequence. Mutagenesis is powerful genetic tool to study and characterize functional elements of gene structure.

3-6.2 Classification

Mutagenesis is classified under different category depending on the target site:

1) Site directed/ site specific mutagenesis- A specific site or stretch of some nucleic acids is altered in a predetermined way by exploiting this approach. This is accomplished by specific nucleic acid deletion or substitution to the targeted site of mutation.
2) Mismatch mutagenesis- Mismatch mutagenesis is applied to create a desired point mutation in a unique site precisely by introducing a mismatched nucleic acid base. This method is generally used to determine the functional groups of a protein encoded by that gene.

Several methods have been developed till now to obtain a successful site directed mutation in the targeted site. Some of them are discussed below-

3-6.2.1 Kunkel's method

In the early 1980's, Thomas Kunkel developed a technique which reduced the need of selection of the mutants. Candidate DNA fragment to be mutated is inserted into a phagemid such as M13mp18/19 and then transformed into an *E. coli* strain deficient in dUTPase(*dut*) and uracil deglycosidase (*ung*). These enzymes are part of a DNA repair pathway that protects the bacterial chromosome from mutations by the spontaneous deamination of dCTP to dUTP. The dUTPase deficiency prevents the breakdown of

dUTP, resulting in increased level of dUTP in the cell. The uracil deglycosidase deficiency prevents the removal of uracil from newly-synthesized DNA. As the double-mutant *E. coli* replicates the phage DNA, its enzymatic machinery may therefore mis-incorporate dUTP instead of dTTP, resulting in single stranded DNA which contains ssUDNA. The ssUDNA is extracted from the bacteriophage released into the medium, and used as a template for mutagenesis. An oligonucleotide containing the desired mutation is used for primer extension. The heteroduplex DNA containing one parental non-mutated strand of dUTP and a mutated strand containing dTTP is then transformed into an *E. coli* strain carrying the wild type *dut* and *ung* genes. Here, the uracil-containing parental DNA strand is degraded, so that nearly all resulting DNA consists of the mutated strand.

3-6.2.2 Cassette mutagenesis

In cassette mutagenesis, a fragment of DNA containing the mutation in the gene of interest is synthesized. Then it has been inserted into a plasmid after the cleavage by a restriction enzyme at a particular site in the plasmid and subsequent ligation. Usually the restriction enzymes that cleave the plasmid and candidate oligonucleotide are same to allow generation of sticky ends of the plasmid and insert to ligate to one another. This method can generate mutants at close to 100% efficiency, but in practice limited by the availability of suitable restriction sites.

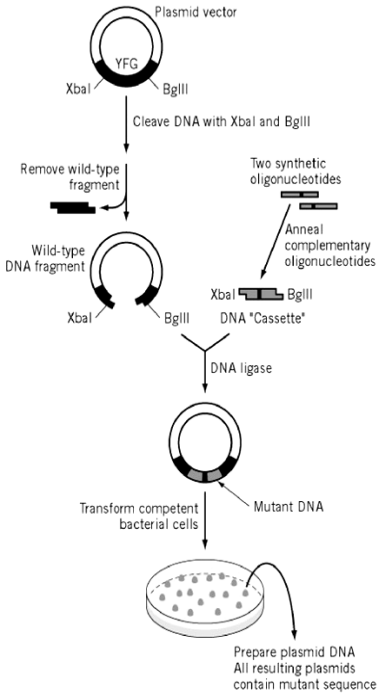


Fig 3-6.2.2 Schematic flow diagram of cassette mutagenesis
(Source: http://what-when-how.com/molecular-biology/site-directed-mutagenesis-part-1-molecular-biology)

A plasmid containing a target gene YFG (black segment) is cleaved with restriction enzymes XbaI and BglII, each of which has unique restriction site in the entire plasmid. The reaction mixture is separated by agarose gel electrophoresis. Two ss oligonucleotides are synthesized by automated DNA synthesis. The sequences of the oligonucleotides are complementary to each other and differ from wild-type sequence at only a single position (black stripe) containing the desired changes. The

oligonucleotides are mixed that helps in hybridization of the two strands by complementation. The ends of the duplex fragment are single-stranded, sticky ends that join with XbaI and BglII sites. The DNA cassette is mixed with the isolated fragment and two molecules are covalently joined by action of T4 DNA ligase. The ligated DNA is transformed into E. coli, and drug-resistant colonies are selected. Plasmid DNA is prepared from individual bacterial colonies. Since the two linear fragments themselves cannot transform E. coli, all colonies contain plasmids with the mutant sequence.

3-6.2.3 PCR based site Directed Mutagenesis

The limitation of restriction sites in cassette mutagenesis may be overcome using polymerase chain reaction with specific oligonucleotide "primers", such that a larger fragment may be generated covering two convenient restriction sites. The exponential amplification in PCR produces a fragment having the desired mutation in sufficient quantity to be isolated from original, unmutated plasmid by gel electrophoresis. There are many variations of the same technique. The simplest method results the mutation site towards one of the ends of the fragment whereby one of two oligonucleotides used for generating the fragment contains the mutation. This involves a single step PCR, but still has the problem of requiring a suitable restriction site near the mutation site unless a very long primer is used. Other variations therefore employ three or four oligionucleotides, two may be non-mutagenic oligonucleotides that cover two convenient restriction sites and generate a fragment that can be digested and ligated. Mutagenic oligonucleotides may be complementary to a location within that fragment well away from any convenient restriction site. These methods require multiple steps of PCR so that the final fragment to be ligated can contain the desired mutation.

3-6.2.3 *In-vivo* site directed mutagenesis

In vivo genome manipulation through site-directed mutagenesis can be done by various methods described as below.

a) *In vivo* site-directed mutagenesis with synthetic oligonucleotides:

This system combines the versatility of synthetic oligonucleotides for targeting with the practicality of a general selection system. It provides for an enormously wide variety of genome modifications via homologous recombination. Exceptional high frequencies of mutations can be obtained when a site-specific double-strand break (DSB) is induced within the locus targeted by the synthetic oligonucleotides. (Storici F *et al, 2006*)

b) Trans-placement "pop-in pop-out"

Most gene targeting experiments have been used to disrupt endogenous loci, resulting in targeted null alleles (strategy is often termed as ‘gene knockout’). Two types of vector have been developed for this purpose: insertion vectors and replacement (or transplacement) vectors.

Insertion vectors are linearized within the homology region, resulting in insertion of entire vector into the target locus. This type of vector disrupts the target gene but leads to a duplication of the sequences adjacent to the selectable marker.

Replacement vectors are designed so that the homology region is collinear with the target. The vector is linearized outside the homology region prior to transfection, thus crossover events in which endogenous DNA is replaced by the insert DNA. With this type of vector, only sequences within the homology region (not the vector backbone) are inserted. Thus for gene knockout, the homology region itself must be interrupted. Insertion and replacement vectors

are equally efficient, but replacement vectors have been used in the majority of knockout experiments.

c) Direct gene deletion and site-specific mutagenesis with PCR and one recyclable marker using long homologous regions.

Several methods have been developed till now to obtain a successful mismatch mutation (transition/ transversion) in the targeted site. Some of them are discussed below:

3-6.2.4 Spontaneous hydrolysis

DNA is not entirely stable in aqueous solution. Under certain physiological conditions glycosidic bond may be hydrolyzed impulsively and thousands of purine sites in DNA are estimated to be depurinated daily in a cell. Several DNA repair pathways exist for the DNA, however, if apurinic site is irrepairable, mis-incorporation of nucleotide may take place during replication. Adenine is incorporated by DNA polymerases in an apurinic site. Cytidine may also deaminated to uridine at lower rate of depurination and can form G to A transition. Eukaryotic cells also contain 5'-methylcytosine, may be involved in the control of gene transcription, which can become deaminated into thymine.

3-6.2.5 Modification of bases

Bases may be modified endogenously by normal cellular molecules. For example DNA may be methylated by S-adenosylmethionine (SAM) and glycosylated by reducing sugars.

Many compounds, such as PAHs, aromatic amines, aflatoxin and pyrrolizidine alkaloids, may form reactive oxygen species (ROS) catalyzed by cytochrome P450. These metabolites form adducts with the DNA, which can cause errors in replication, and thus bulky aromatic adducts may form stable intercalation between bases and block replication. The adducts may also induce conformational changes in the DNA. Some adducts also result in the depurination of the DNA, however it may vary how significant such depurination is in generating mutation.

Alkylation and arylation of bases can cause errors in replication. Some alkylating agents such as N-Nitrosamines may require the catalytic reaction of cytochrome-P450 for the formation of a reactive alkyl cation. N7 and O6 of guanine and the N3 and N7 of adenine are most susceptible to attack. N7 of guanine adducts form the bulk of DNA adducts, but they appear to be non-mutagenic. Alkylation at O6 of guanine however is harmful because excision repair of O6-adduct of guanine may be poor in some tissues such as the brain. The O6 methylation of guanine can result in G to A transition, while O4-methylthymine can be mis-paired with guanine. The type of the mutation generated however may be dependent on the size and type of the adduct as well as the DNA

sequence. Ionizing radiations and reactive oxygen species often oxidize guanine to produce 8-oxoguanine.

3-6.2.6 Cross linking of DNA

Some alkylating agents produce cross linking of DNA. Some natural occurring chemicals such as psoralens after activation by UV radiation, and nitrous acid may also promote cross linking. Inter-strand cross-linking is more lethal as it blocks replication and transcription and can cause chromosomal breakages and rearrangements. Some cross linkers such as cyclophosphamide, mitomycin C and cisplatin are used as anticancer agent because of their high degree of toxicity to proliferating cells.

3-6.2.7 Dimerization

UV radiation promotes the formation of a cyclobutyl ring between adjacent thymines, resulting in the formation of pyrimidine dimers. In human skin cells, thousands of dimers may be formed in a day due to normal exposure to sunlight. Human DNA polymerase kappa help bypass these lesions in an error-free manner however, individuals with defective DNA repair function, such as sufferers of *Xeroderma pigmentosum* (autosomal recessive genetic disorder), are sensitive to sunlight and may be prone to skin cancer.

3-6.2.8 Intercalation between bases

The planar structure of chemicals such as ethidium bromide and proflavine allows them to insert between stacked bases in ds DNA. This insert causes the DNA's backbone to stretch and makes slippage during replication more likely to occur since the bonding between the strands is made less stable by the stretching. Forward slippage results in deletion mutation, while reverse slippage results in an insertion mutation. The intercalation of anthracyclines such as daunorubicin and doxorubicin also interferes with the functioning of the enzyme topoisomerase II, blocking replication as well as causing mitotic homologous recombination.

3-6.2.9 Backbone damage

Ionizing radiations may produce highly reactive free radicals that can break bonds in the DNA. Double-stranded breakages are particularly damaging and tough to repair, forms translocation and deletion of part of a chromosome. Alkylating agents like mustard gas cause breakages in the DNA backbone. Oxidative stress produces highly reactive oxygen species (ROS) which cause damage to DNA. Incorrect repair of other damages induced by the highly reactive species also lead to mutations.

3-6.2.10 T-DNA mediated mutagenesis:

Transfer-DNA (T-DNA) insertion is a highly effective mutagen for genome-wide mutagenesis (Krysan *et al.* 1999). It has been widely exploited to produce insertion mutants in *Arabidopsis thaliana* for functional characterization of the every gene in the genome. Even if mRNA is transcribed, T-DNA sequence may contain stop codons, resulting in early translation termination (Krysan *et al.* 1999). Although it is not an ideal technique, T-DNA insertion mutagenesis has been a powerful technique to link genes to phenotypes

For plant system, homologous recombination machinery is not very effective. Moreover, active transposons do not exist in every species including *Arabidopsis*. Under such limitations, T-DNA mediated insertional mutagenesis provides a suitable tool for insertional mutagenesis. T-DNA is the transferred DNA from tumour inducing plasmid from *Agrobacterium sp.* T-DNA is widely used for successful transformation in plants. During T-DNA mediated transformation, the mutated locus also gets tagged which helps further isolation of the mutated gene.

3-6.2.11 Transposon mediated mutagenesis:

Transposons are DNA elements which have ability to move from one region to another region in the genome. The shifting of transposon element from a heterochromatin region to a transcriptionally active region results in insertional mutagenesis. It can be exploited as a genetic tool for analysis of gene and protein function. The study of transposons is well described in the case of Drosophila (in which P elements are most

commonly used) and in Thale cress (*Arabidopsis thaliana*) and bacteria such as *Escherichia coli.*

Transposon based mutagenesis is a biological process that helps in genes to be transferred to a host organism's chromosome, modifying the function of an extant gene on the chromosome and causing mutation. Transposon mutagenesis was first studied by *Barbara McClintock* in the mid-20th century with corn for which she was awarded Nobel prize also.

In case of bacteria, transposition mutagenesis is usually done by the help of a plasmid from which a transposon is extracted and inserted into the host chromosome. This usually requires a set of enzymes including transposase to be translated.

Currently transposons can be used in genetic research and recombinant genetic engineering for insertional mutagenesis. Insertional mutagenesis is when transposons act like vectors to help remove and integrate genetic sequences. Due to their simple and intricate design and inherent ability to move DNA sequences, transposons are highly compatible for transducing genetic material, enable them as an ideal genetic tool.

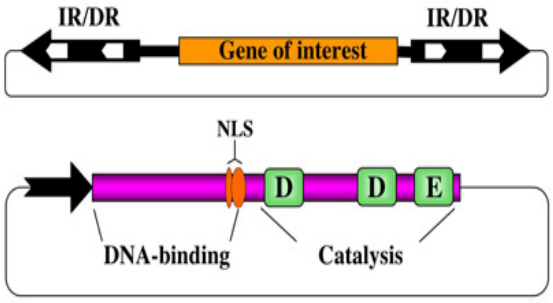


Fig 3-6.2.11 [A]

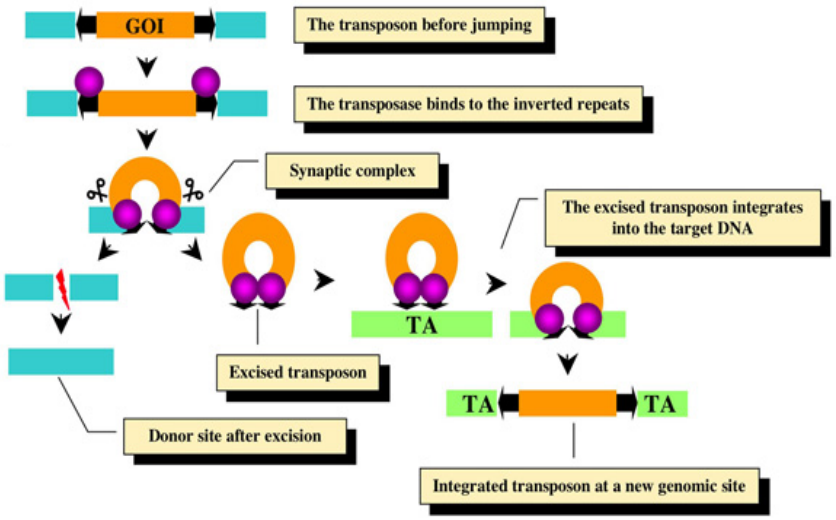


Fig 3-6.2.11 [B]

Fig 3-6.2.11 The *Sleeping Beauty* transposon system.

(Source: https://www.mdc-berlin.de/8201913/en/research/research_teams/mobile_dna/research)

Figure [A] describes the components and structure of a two-component gene transfer system based on Sleeping Beauty. A gene of interest (orange box) to be mobilized is cloned between the terminal inverted repeats (IR/DR, black arrows) that contain binding sites for the transposase (white arrows). The transposase gene (purple box) is physically separated from the IR/DRs, and is expressed in cells from a suitable promoter (black arrow). The transposase consists of an N-terminal DNA-binding domain, a nuclear localization signal (NLS) and a catalytic domain characterized by the DDE signature.

Fig [B] illustrates the mechanism of Sleeping Beauty transposition system. The transposable element carrying a gene of interest (GOI, orange box) is maintained and delivered as part of a DNA vector (blue DNA). The transposase (purple circle) binds to its sites within the transposon inverted repeats (black arrows). Excision takes place in a synaptic complex. Excision separates the transposon from the donor DNA, and the double-strand DNA breaks that are generated during this process are repaired by host factors. The excised element integrates into a TA site in the target DNA (green DNA) that will be duplicated and will be flanking the newly integrated transposon.

Bibliography

Krysan P J, Young JC, Sussman MR (1999) T-DNA as an Insertional Mutagen in Arabidopsis. Plant Cell 11:2283-2290

Kunkel TA.1985. "Rapid and efficient site-specific mutagenesis without phenotypic selection." *Proceedings of the National Academy of Sciences U S A.* 82 (2): 488–92

Li Hong Wang.2008.How effective is T-DNA insertional mutagenesis in *Arabodopsis*? J Biochem Tech 1(1):11-20.

Rashtchian A,Thornton C.G. and Heidecker G. 1992. A novel method for site directed mutagenesis using PCR and uracil DNA glycosylase; Genome Res., 2: 124-130.

Stanford *et al.*2001. Gene-trap mutagenesis: past, present and beyond. Nature Reviews Genetics 2, 756-768.

Storici F, Resnick MA. 2006. The delitto perfetto approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. Methods Enzymol.; 409: 329-45.

William L.Stanford, Jason B. Cohn and Sabine P. Cordes.2001.Gene trap mutagenesis: Past, present and beyond; Nat Rev, Vol2.