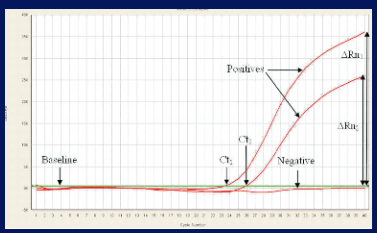
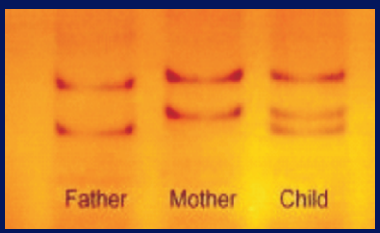
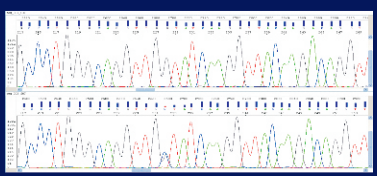
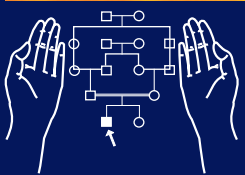
MANUAL OF PCR
IN DIAGNOSTIC PATHOLOGY

Maj Gen Suhaib Ahmed, HI (M)
MBBS; FCPS (Pak); PhD (London)









GRC
Publication

Manual of PCR in Diagnostic Pathology

Maj Gen Suhaib Ahmed, HI (M)
MBBS; FCPS (Pak); PhD (London)

Manual of PCR in Diagnostic Pathology

ISBN - 978-969-9829-00-0

First Edition 2013

Copyright © GRC Publications, 2013
All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher.

Genetics Resource Centre (GRC)
House # 14, Street 16, Race Course Road,
Westridge, Rawalpindi - Pakistan
Tel: 0092-51-2582607
E-mail: info@grcpk.com
Web: www.grcpk.com

Published by:
SAAD PrintoPack
12-Khurshid Palace, Kashmir Road, Rawalpindi Cantt - 46000, Pakistan.
Phone: +92 51 5522711 - 5522811 - Fax: +92 51 5511170
Email: saadprintopack@gmail.com

Table of Contents

Section-I

Section-II

Preface

Polymerase Chain Reaction (PCR) has made the analysis of genes simple and cheap. A large number of diseases with a genetic basis can be diagnosed by PCR. Prenatal diagnosis of inherited disorders in early pregnancy by PCR has completely changed the concept of management of common inherited disorders like thalassaemia. PCR has also become an essential tool for identification of several infectious diseases. The hazardous procedure of viral culture is almost completely replaced by simple accurate and rapid PCR based assays. Real time PCR is being used to precisely quantify the target DNA or RNA of infectious agents or a clone of malignant cells.

There is no shortage of books on PCR. But most such books are written by scientists who have relatively limited knowledge of the clinical applications of PCR. Moreover the books on PCR mostly focus on a limited range of applications. The residents in Pathology and laboratory technologists have always looked for a book that gives an easily comprehensible concept of PCR applications in diagnostic Pathology. While this book would be useful for the pathologists and lab technicians, many scientists researchers and university students may also benefit from it.

The book has two sections. The first section deals with the basic technique of PCR and related procedures like nucleic acid extraction and post PCR analysis. The second section of the book includes common applications of PCR in inherited, neoplastic and infectious diseases. In addition chapters on forensic DNA testing and study of donor chimerism are also included.

I would be grateful if the readers could point out the mistakes, errors or omissions in the book. This feedback will help in improving the standard of the future editions of the book.

Maj Gen Suhaib Ahmed, HI (M)
Rawalpindi
Jan 2013

Section-I

1. The Genome

The human genome comprises DNA sequences mostly contained in the nucleus. A small portion of DNA is also present in the mitochondria. The nuclear DNA is present in chromosomes. Each chromosome is a single coiled up molecule of DNA and a supporting protein backbone. The human genome comprises 23 pairs of chromosomes including 22 autosomes and one pair of sex chromosomes.

Deoxyribonucleic acid (DNA)

DNA is a double stranded molecule in which each strand is made of individual units called nucleotides. The nucleotides consist of a deoxyribose (5-carbon sugar), a nitrogen containing base attached to the sugar, and a phosphate group (Fig. 1.1). There are four different nucleotides each differing by the nitrogenous base. The bases are either purines (adenine and guanine) or pyrimidines (cytosine and thymine). The four bases are abbreviated as A, G, C and T. The deoxyribose sugar has 5 carbon atoms numbered 1', 2', 3', 4', and 5'. A hydroxyl group on the 5' and 3' carbons is attached to a phosphate group to form the DNA backbone. The 5' end of the strand is the starting point of DNA molecule while the 3' end is the finishing point. During the synthesis of a new strand nucleotides are added at the 3' end and the strand elongates from 5' to the 3' end.

The two strands of DNA run in opposite directions and wind around each other to form a right handed spiral. The strand that runs from left to right (➔) is called the forward strand and the strand running from right to left (➔) is called the reverse strand (Fig 1.2). The nitrogenous bases of the nucleotides face towards the interior of the helix. The two strands are kept together through hydrogen bonds between purines and pyrimidines. Adenine (A) forms two hydrogen bonds with thymine (T) and cytosine (C) forms three hydrogen bonds with guanine (G) on the opposite strands. The G-C bond is stronger than the A-T bond. This makes the G-C rich areas of DNA more stable than the rest. The synthesis of DNA starts with separation of the two strands followed by addition of nucleotides on each strand from 5' end to the 3' end.

DNA sequence

The nucleotide composition of a DNA molecule is called its “sequence”. By convention and for convenience the DNA sequence of only the forward strand is written (Fig 1.3). The sequence is written from left to right starting from the 5' end and going towards the 3' end. Only the abbreviations of the nucleotides i.e. G, A, T and C are used. Occasionally other abbreviations are also used e.g. Y for pyrimidines, R for purines, and N for any nucleotide.

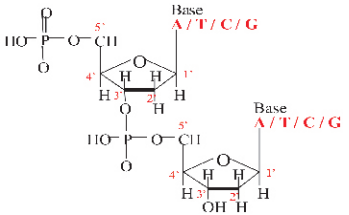


Fig 1.1. The chemical structure of DNA molecule.

```
Forward strand →
5′-CTGGACTTGACTGATACGTGCATTAGCATGTCCCAGTGCTAGCTAGATCAT-3′
3′-GACCTGAACTGACTATGCACGTAATCGTACAGGGTCACGATCGATCTAGTA-5′
                                        ← Reverse strand
```

Fig 1.2. Double stranded DNA. The forward and the reverse strands run in opposite directions.

```
5′-CGGAGCTGCGCGATGGCGAACTCAAGGAGCACATCAGCCGCGTCCACGCCGCC
AACTACGGTGTTTACGGTGCCCGCAAAGTGTGGCTAACCCTGAACCGTGAGGCATC
GAGGTGGCCAGATGCACCGTCGAACGGCTGATGACCAAACTCGGCCTGTCCGGGAC
CACCCGCGGCAAAGCCCGCAGGACCACGATCGCTGATCCGGCCACAGCCCGTCCCG
CCGATCTCGTCCAGCGCCGCTTCGGACCACCAGCACCTAACCGGCTGTGGGTAG-3′
```

Fig. 1.3. Mycobacterium tuberculosis insertion sequence (IS6110) (GenBank accession: AE000516.2).

Ribonucleic Acid (RNA)

RNA in a cell is seen as messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA). The basic structure of all types of RNA molecules is similar. It exists as single stranded molecule. The components of RNA are the same as of DNA with two major differences i.e. uracil (U) replaces thymine (T) and ribose sugar replaces deoxyribose. Uracil is a pyrimidine that is structurally similar to thymine, and can base-pair with adenine.

DNA and RNA secondary structures

Three dimensional shape of a segment of single stranded DNA or RNA is often called "secondary structure". It is formed by hydrogen bonding between the adjacent nitrogenous bases. DNA is a double stranded molecule; however, during in-vitro experiments single stranded DNA is often encountered. PCR primers are a typical example. Single stranded DNA and RNA may give rise to helices, loops and hairpin structures (Fig. 1.4). DNA strands rich in G-C content are especially prone to develop secondary structures.

```
┌─GTGTCT-5′
A   |||                    5′-TCTGTGATGCAGCATCGTA-3′
└─TGCAGCATCGTA-3′
```

Fig 1.4 Example of hairpin secondary structure of a short sequence of DNA.

The genome resources

DNA analysis like Polymerase Chain Reaction (PCR) requires a basic knowledge of the sequence to be analyzed. The DNA sequence of interest can be obtained from a number of sources. The lack of knowledge about such resources is an important reason for underutilization of PCR in diagnostic pathology.

The most common source of genomic information is the National Centre for Biotechnology Information (NCBI) website. This may be accessed at the web address http://www.ncbi.nlm.nih.gov/Database/index.html

Information at the NCBI website is contained in databases including PubMed, GenBank (nucleotide and protein sequences), protein structures, complete genomes, taxonomy and others. Searching of the major databases is done through “Entrez” which is an integrated, text-based search engine and retrieval system.

GenBank® is a genetic sequence database at the National Institute of Health (NIH), Bethesda USA. It is an annotated collection of all publicly available DNA sequences. As of August 2009 there are approximately 106,533,156,756 bases in 108,431,692 sequence records in the traditional GenBank divisions and 148,165,117,763 bases in 48,443,067 sequence records in the Whole Genome Shotgun (WGS) projects division. The GenBank contains data of over 300,000 species including 1000 complete genomes of bacteria and archaea.

The information is submitted to GenBank by individual researchers and project groups all over the world. The database is increasing by about 1700 specie records every month. About 12% of the sequences in the GenBank are of human origin. The files in the GenBank are sorted into 'divisions' such as bacteria (BCT), viruses (VRL), primates (PRI) and rodents (ROD) etc.

The complete release notes for the current version of GenBank are available at the NCBI website. A new release is made every two months. GenBank is part of the International Nucleotide Sequence Database Collaboration which includes the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI. These three organizations exchange data on a daily basis.

When a new sequence is received the GenBank staff assigns it a unique accession number. Each sequence is also given a unique NCBI “gi” identifier number. A third identifier is the version of the sequence. The older as well as the newer versions of the sequences are kept in the database. For example complete nucleotide sequence of the human beta globin complex can be accessed by entering the accession no NG_000007.3 or the gi number 28380636 in the Entrez life sciences search engine. The version of the sequence, if any, is identified by the digit after a decimal in the accession number. In the accession number NG_000007.3, 3 indicates the version. The sequence records can be viewed in the GenBank, FASTA or Graphic format. The nucleotide sequence can be downloaded in the FASTA format which is the most convenient format for use in primer designing etc.

How to get the accession number of the sequence of interest?

The accession number of a sequence is usually obtained from the scientific publications. Since GenBank database is also linked to the scientific literature via PubMed and PubMed Central a sequence can also be found by open search through the Entrez. For example entering "human beta globin gene" in the Entrez would display the related links.

Basic Local Alignment Search Tool (BLAST)

BLAST is a fundamental and most frequently performed function of the GenBank. It is a family of search tools that allows comparison and finding similarities between various sequences. BLAST allows comparison of a sequence data with any of the known sequences available in the GenBank. For example the primer BLAST feature allows searching of the entire human, bacterial, viral or parasitic genomes to know whether a newly designed primer would cross anneal with a homologus sequence elsewhere or not.

Searching the human β-globin gene (GenBank NG_000007.3)

1. Enter the accession number NG _000007.3 in the and Entrez website press go (Fig. 1.5)
2. Click on the area pointed out by red arrow in Fig 1.5. This will display the complete record of the desired sequence in a new window (Fig 1.6).
3. The sequence can be displayed in the FASTA or the graphic format by clicking on the appropriate area (red arrow in Fig 1.6).
   The sequence is opened in FASTA format in another window (Fig. 1.7).
   The sequence may be cut and pasted from this window.

Fig. 1.5. Entrez provides a search engine for exploring the NCBI website. Any nucleotide sequence record may be accessed by entering the GenBank accession number. All related entries like nucleotide sequence, PubMed citations etc. can be accessed by clicking on the related links.

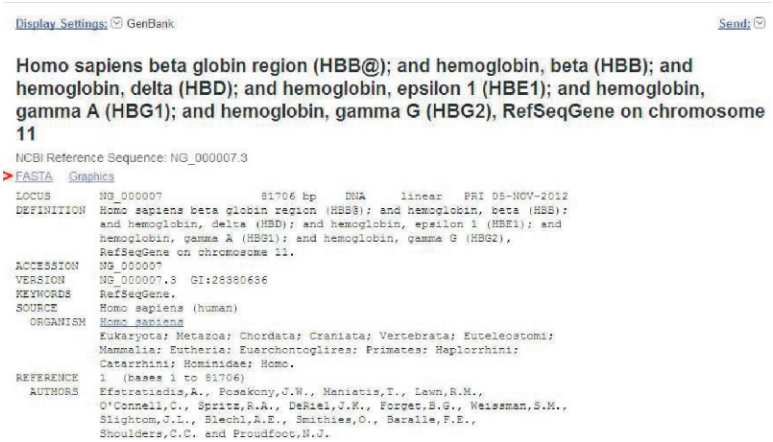


Fig. 1.6. NCBI Reference sequence of human beta globin region. The sequence record can be opened in graphic or FASTA format by clicking on the related links (red arrow).

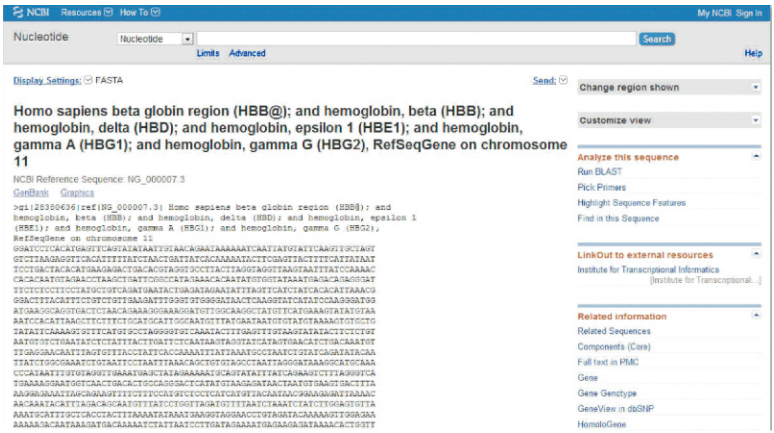


Fig. 1.7. Sequence of the human beta globin region in FASTA format. The sequence can be cut and pasted anywhere for further use.

Bibliography

1. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J Sayers EW (2010) GenBank. Nucleic Acids Res 38 (Database issue): 46-51.

2. Extraction of Nucleic Acids

DNA Extraction

The first step in DNA analysis is to get a good quality sample; a process commonly known as "extraction". Since DNA is a very long molecule it can easily get broken by vigorous shaking during the process of extraction. Therefore gentle and careful handling in the processing is essential. DNA can also be destroyed by DNase that is commonly present in the environment or in the bacteria that may contaminate the sample.

There are three basic steps in a DNA extraction:

1. Removing the membrane lipids by detergents to expose DNA in the nucleus of the cell.
2. Removal of proteins by protease digestion and subsequent precipitation by phenol or other agents.
3. Precipitation of DNA with ethanol or isopropanol.

Sources of DNA

DNA can be extracted from any source that contains nucleated cells. It is most commonly extracted from blood collected in EDTA. EDTA has an additional advantage of chelating magnesium ions which reduce the activity of DNA degrading enzymes of cellular or bacterial origin. Heparin is not a good anticoagulant as it interferes with subsequent PCR. The blood may be kept at 4°C for a few days without causing any significant loss in the yield of DNA. Blood contaminated with bacteria becomes unsuitable for DNA extraction. The blood sample may be kept frozen for long periods before DNA extraction. DNA can also be extracted from bone marrow aspirates or bone marrow smears on slides. Archival bone marrow slides stored at room temperature for several years have been used to extract good quality DNA.

Buccal smear on cotton swab or mouth wash is another easily available source of DNA. This is especially useful for field work. Solid fresh tissues, like surgical biopsy specimens, chorionic villi and tissues collected at autopsy are also used for DNA extraction. DNA can also be extracted from hair root, blood stains, archival bones etc. Fixation of the tissue with formaline can make DNA extraction very difficult. Special processing protocols may be required to extract DNA from paraffin embedded tissues.

Choice of the method

The standard method of DNA extraction uses phenol chloroform for protein precipitation. Keeping in view the toxicity of phenol, methods have been developed to precipitate proteins without using phenol. A large number of commercial kits are also available that are time as well as cost effective. Some of the methods can also be automated for large scale DNA extraction. A quick

method of DNA extraction is by ion exchange resin Chelex 100.
In the subsequent section phenol chloroform method is described in detail. It is robust and cost effective and consistently gives good quality high molecular weight DNA.

DNA extraction from whole blood

1. In a 5ml conical tube take 3 ml blood in EDTA.
2. Centrifuge at 3000 rpm for 5-10 minutes.
3. Remove the supernatant plasma leaving behind the buffy coat and the RBCs.
4. Add red cell lysis buffer (Table 2.1) 2-3 times the volume of the red cells and mix by inverting a few times.
5. Centrifuge at 3000 rpm for 5-10 minutes and discard the supernatant.
6. Repeat the above step once more if the cell pellet contains too many RBCs.
7. Add 0.7ml cell lysis buffer (Table 2.1) and transfer the contents to a 1.5ml Eppendorf tube.
8. Add 20µl Proteinase-K (Table 2.1) and mix by gentle vortexing.
9. Incubate at 37°C overnight or at 56°C for two hours.
10. An alternate to the standard cell lysis buffer is the lysis buffer containing guanidine isothiocyanate (Table 2.1). The later can effectively breakdown proteins without adding proteinase-K. Use 0.7ml cell lysis buffer with guanidine to lyse the WBC pallet. Shake well on vortex and place at 37°C overnight.
11. Add 250µl buffered phenol (Table 2.1) and 250µl chloroform.
12. Vortex for a few seconds.
13. Centrifuge at 10,000 rpm for 2 minutes in a micro-centrifuge.
14. Carefully remove the subnatent phenol layer with a pasture pipette leaving behind the clear watery supernatant.
15. Repeat the above step if cloudiness still remains in the supernatant.
16. Add 500µl chloroform and vortex for a few seconds.
17. Centrifuge at 10,000 rpm for 2 minutes.
18. Remove as much of the subnatent chloroform as is possible leaving behind clear supernatant DNA solution.
19. Add 150µl 7.4M ammonium acetate solution.
20. At this stage the DNA solution left from the previous step should be approximately 500µl. Fill the Eppendorf tube to its top with pure ethanol (about 1ml). This will make a final concentration of 70% ethanol in which the DNA forms a whitish precipitate.
21. Gently invert the Eppendorf tube 3-4 times and watch for a whitish hairball like precipitate of DNA.
22. Centrifuge at 10,000 rpm for 2 minutes.
23. Carefully remove ethanol by inverting the tube and leaving behind the DNA pellet at the bottom of the tube.
24. Add 500µl of fresh 70% ethanol and gently vortex to give the DNA pellet a good wash in ethanol.
25. Centrifuge at 10,000 rpm for 1 minute.

26. Remove as much of the top ethanol layer as is possible. Leaving some ethanol behind can interfere with subsequent DNA hydration.
27. Dry the DNA pallet for 5-10 minutes by keeping the tube in inverted position on a clean tissue paper.
28. Dissolve the DNA in DNase free water. Ordinary distilled water can also be used if DNase free water is not available. The amount of water to be added depends on the yield as seen in the DNA pellet. On an average the DNA extracted from 3ml of blood with a normal white cell count can be dissolved in 200-300µl of water to give a final concentration of ~200ng/µl. It is safe to add less water as the concentrated solution can be diluted further whereas a diluted DNA can not be concentrated if required in future!
29. Leave the DNA solution at 37°C for 15-30 minutes.
30. DNA may be stored at 4°C for a few weeks, at -20°C for several months and at -80°C for several years.
31. The DNA solution is fairly stable at room temperature for many days. It can be transported from one place to another without being kept in ice.

DNA extraction from CVS and fresh tissues

1. Take approximately 25-50mg of fresh tissue (chorionic villi, skin, or other solid tissues) in 0.5ml of cell lysis buffer (Table 2.1).
2. Add 20-40µl of Proteinase-K (Table 2.1) depending on the amount of tissue.
3. Keep at 37°C overnight. Allow longer incubation or add more Proteinase-K if the tissue is not completely digested/dissolved.
4. Proceed as step 11 onwards of the DNA extraction protocol.

DNA extraction from archival bone marrow slides

1. Take a slide of bone marrow smear that has good number of cells.
2. Layer about 0.7ml cell lysis buffer (Table 2.1) on the smear.
3. Gently scratch the smear from the slide with a wooden stick and transfer the contents to an Eppendorf tube.
4. Add 20µl Proteinase-K (Table 2.1) and keep at 37°C overnight.
5. Cell lysis buffer with guanidine (Table 2.1) may be used instead of the standard cell lysis buffer and Proteinase-K.
6. Proceed as step 11 onwards of the DNA extraction protocol.

Chelex method of DNA extraction

1. Make 5-7% solution of chelex, aliquot 300µl in 1.5ml Eppendorf tubes and refrigerate.
2. Take 300µl blood and add 3ml distilled water or RBC lysing solution (Table 2.1) to lyse the red cells.
3. Centrifuge at 5000 rpm for 2 minutes to pallet the white cells.
4. Repeat red cell lysis step if the white cell pallet contains too many red cells.
5. Add 300µl 5-7% chelex solution to the white cell pallet and vortex for 15-20 seconds.

6. Place the tube in a heating block at 95°C for 20 minutes.
7. Vortex for 15-20 seconds.
8. Centrifuge at 10,000 rpm for 2 minutes.
9. Transfer the supernatant to a fresh Eppendorf tube and use as source of DNA.
10. The DNA extracted by Chelex method may contain some residual haemoglobin especially when the white cell pallet contains red cells. Such DNA may give excessive background fluorescence in real time PCR applications.

DNA Extraction from archival bones

Extraction of DNA from dried bones is always challenging. The source of DNA in a bone sample lies in the osteocytes located in the dense cortex. Therefore it is essential to free the osteocytes for penetration by the cell lysing reagents. The dense cortical bone is first converted to fine powder by using a file or a saw. The archival bones recovered from graves etc. are also heavily contaminated by dust and other PCR inhibitors. Therefore the bones must first be cleaned to remove any possible contaminants.

The following protocol gives reasonably good results.

1. Wash the bone surface with 0.5-1M EDTA and rinse in distilled water. Dry the bone in air and choose a thick cortical portion from a long bone like femur. Spongy soft bones usually harbour contaminants and should be avoided for DNA extraction.
2. Convert the cortical bone to fine powder by using a saw or a file. The process of filing is done carefully to avoid generation of heat that may degrade DNA. Usually 5-10gm of bone powder is enough for processing.
3. The next step is to decalcify the bone particles by adding sufficient 0.5M EDTA to the bone powder and leaving at room temperature for 48 hours. Frequent agitation or vortexing helps this process.
4. Centrifuge for 3 minutes at 13000 rpm to remove the supernatant EDTA solution.
5. Wash the bone particles twice in distilled water.
6. Add 0.5ml lysis buffer (Proteinase K 20mg/ml, 10µl 1M Tris-HCl, 2µl 0.5M EDTA, 100µl 10% SDS and 200µl distilled water). Incubate at 56°C overnight. If bone particles are not completely dissolved the step may be repeated with addition of fresh Proteinase K until the bone particles are completely dissolved. This may take 2-3 days.
7. The further steps of DNA extraction of DNA from the dissolved bone are the same as described in the section of DNA extraction from blood by phenol chloroform method.

DNA extraction from paraffin embedded tissues

Extraction of DNA from paraffin embedded tissue blocks that are fixed in formaline is difficult. DNA can be extracted by taking 2-3 microtome sections in xylene to dissolve the wax. The tissue is air dried and can be processed as for the

fresh tissue or the chelex protocol described above.
Many commercial kits are also available that give consistently good quality results.

Measurement of DNA concentration

Most PCR applications work well at DNA concentration of 100-300ng/µl. This concentration can be achieved by following the guidelines given in the extraction protocol. However, in applications using genetic analyzer it becomes very critical to know the exact concentration of DNA. Several methods are available to know the DNA concentration and its purity.

Optical density (OD) method

DNA and RNA absorb UV light at 260nm. The OD of DNA solution measured at 260nm can be used to calculate the concentration of DNA or RNA. The following example can be used to calculate DNA in an unknown solution:

Make 1: 100 dilution of DNA in distilled water (20µl + 2ml)
Take OD at 260nm
DNA concentration (ng/µl) = 50 x dilution factor x OD
Example:
OD at 260nm: 0.068
Concentration: 50 x 100 x 0.068 = 340 ng/µl or 0.340 µg/ml

The optical density method may also be used to determine the protein content of DNA. Proteins left over from the extraction procedure can interfere in PCR and therefore it sometimes is required to know the purity of the extracted DNA. Proteins absorb UV light at 280nm. In a good DNA sample the ratio of OD at 260nm and 280nm should be above 1.8. Ratio below 1.8 indicates protein contamination in the DNA solution.

Fluorometry

Commercial kits based on fluorescent dyes like SYBR Green can be used for DNA quantitation. The fluorescence given by standards of known DNA concentration is used to know the concentration of an unknown sample of DNA.

Gel electrophoresis

Gel electrophoresis may be used to compare the quantity of DNA in a sample with a known standard of DNA. The unknown DNA is amplified by PCR. A DNA of unknown concentration is amplified and it is run on a gel along with a commercially available 100bp ladder. The intensity of the amplified DNA bands of the unknown sample is compared with that of the bands of the allelic ladder. An approximate estimate of the DNA quantity in the unknown sample is made by comparison with the bands of the known concentration of DNA in the allelic ladder. An additional advantage of the method is that the quality of the unknown sample can also be judged. A sample with fragmented DNA would give a uniform smearing effect in the lane (Fig 2.1).

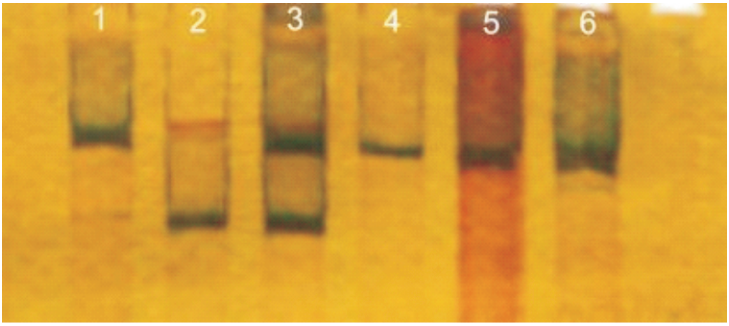


Fig. 2.1 Assessment of DNA quality by gel electrophoresis of amplified DNA. The sample in lane 5 shows fragmented DNA that appears as dark brown smearing in the background of the amplified DNA band. The samples in lane 1, 3 and 6 also show increase in the background but this is due to an unwanted high concentration of DNA resulting in over amplified PCR products. The later can be improved by diluting the DNA. Lane 4 shows the best result with a clean amplification product (sharp band) and almost no background.

Real time PCR

Accurate DNA quantification can also be done by real time PCR. The unknown sample is amplified along with serial dilutions of a DNA standard with known concentration. The results of Ct values are plotted against various DNA concentrations of the standard and the unknown (Chapter 6).

RNA Extraction

RNA is easily degraded by RNA digesting enzymes present in the environment. Since RNA isolation procedure takes place in a strong denaturant that renders RNase inactive, the integrity of RNA is mostly at risk either before or after the extraction. Therefore handling of the sample prior to extraction and storage of RNA after extraction are extremely critical.

A single step RNA extraction reagent is commercially available as TRI reagent. It is a phenol-based reagent that contains a combination of denaturants and RNase inhibitors. The RNA is separated by centrifugation from DNA after extraction with an acidic solution containing guanidinium thiocyanate, sodium acetate, phenol and chloroform. The total RNA remains in the upper aqueous phase, while most of DNA and proteins remain either in the inter-phase or in the lower organic phase. Total RNA is then recovered by precipitation with isopropanol. The extracted RNA is re-suspended and stored in an RNase free solution.

RNA extraction by TRIzol® Reagent

1. Mix 0.75 ml of TRIzol® Reagent (Invitrogen, USA) with 0.25 ml of sample and lyse cells (or cellular debris) suspended in the sample by passing the suspension several times through a pipette. Use at least 0.75 ml of TRIzol® Reagent per 5-10 x 10^6 cells.

2. If the sample volume is < 0.25 ml, adjust the volume to 0.25 ml with water. The volume ratio of TRIzol® Reagent to sample should be 3:1.
3. Keep the lysate/homogenate for 5 minutes at room temperature to permit complete dissociation of nucleoprotein complexes. At this stage the samples can be stored at -70°C for at least one month.
4. Add 0.2 ml chloroform (free of isoamyl alcohol or any other additive) per 0.75 ml of TRIzol® Reagent.
5. Cover the samples tightly and shake vigorously for 15 seconds.
6. Keep the mixture at room temperature for 2-15 minutes depending on the number of cells in the sample.
7. Centrifuge the mixture at 12,000 g for 15 minutes at 4-10°C. Centrifugation at higher temperature may result in DNA contamination of the aqueous RNA phase making it unsuitable for PCR.
8. The mixture separates into a lower red phenol-chloroform phase, interphase and a colourless upper aqueous phase. RNA remains exclusively in the aqueous phase whereas DNA and proteins are in the interphase and organic phase respectively. The volume of the aqueous phase is about 70% of the volume of TRIzol® Reagent used for homogenization.
9. Transfer the aqueous phase to a fresh tube. Interphase and organic phase may be used for subsequent isolation of DNA and proteins.
10. Precipitate RNA from the aqueous phase by mixing with isopropanol. Use 0.5 ml of isopropanol per 0.75 ml of TRIzol® Reagent used for the initial homogenization.
11. Keep at room temperature for 5-10 minutes and centrifuge at 12,000 g for 8 minutes at 4-25°C. RNA precipitate (often invisible before centrifugation) forms a gel-like or white pellet on the side and bottom of the tube.
12. Discard the supernatant and wash the RNA pellet with 1 ml 75% ethanol. The RNA precipitate can be stored in 75% ethanol at 4°C for one week or -20°C for up to one year.
13. Vortex and centrifuge at 12,000 g for 5 minutes at 4-25°C.
14. Discard the ethanol and dry the RNA pellet for 3-5 minutes at room temperature. It is important not to completely dry the RNA pellet as this will greatly decrease its solubility.
15. Dissolve RNA in DEPC treated water or 0.5% SDS by passing the solution a few times through a pipette tip and incubating for 10-15 minutes at 55-60°C.
16. The total RNA is essentially free of DNA and proteins and should have a 260/280 ratio of 1.6-1.9.
17. Hands and dust are a major source of the RNase contamination. Use gloves and keep tubes closed throughout the procedure.

Extraction of viral RNA

Viral RNA is best extracted by commercial kits that use either silica based purification columns or magnetic beads.

Table. 2.1. Reagents used in DNA extraction

Red cell lysis buffer	
Sucrose:	109.5 g
Tris (pH 7.6):	1.58g
$Mg\ Cl_2$:	476 mg
Triton-X:	10 ml
Sodium azide:	200 mg
Distilled water:	up to 1L
Cell lysis buffer	
Tris (pH 8.0)	7.85 g
Disodium EDTA:	6.68 g
SDS:	20 g
Distilled water:	up to 1L
Cell lysis buffer with guanidine	
Guanidine isothiocyanate:	50 gm
SDS	2 gm
IM Sodium citrate (ph 7.0)	2.5 ml
2-Mercaptoethanol:	0.7 ml
Distilled water:	up to 100 ml
Buffered Phenol	
Phenol	250 g
Distilled water	40 ml
Place at 65° C for 1-2 hrs	
Cool and add 300 mg 8-hydroxyquinoline	
Equilibrate by mixing with equal volume of 1M Tris buffer (pH 8.0)	
Carefully remove the supernatant after allowing phenol to settle down	
Repeat twice equiliration with 1M Tris	
Add 0.4 ml 2-merceptoethanol	
Add 100 ml of 0.1M Tris buffer (ph 8.0)	
Store at 4° C in a dark bottlle	
Buffered Phenol	
Proteinase K:	20 g
Distilled water:	1 ml
Make aliquot of 0.5 ml and store at - 20°C	20°C
Use 20 µl/extraction	
7.4 M Ammonium Actetate	
Ammonium acetate:	57 g
Distilled water:	up to 100 ml

Bibliography

1. Chomczynski P (1993) A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. BioTechniques 15: 532-537.
2. Kleines M, Schellenberg K, Ritter K (2003) Efficient extraction of viral DNA and viral RNA by the chemagic viral DNA/RNA kit allows sensitive detection of Cytomegalovirus, Hepatitis B virus, and Hepatitis G virus by PCR. J Clin Microbiol 41: 5273–5276.
3. Green MR, Sambrook J (2012) Molecular Cloning: A Laboratory Manual. (Fourth Edition). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

3. Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) is a technique that is used for amplification of DNA. The double stranded DNA is denatured by heating at ~94°C. As a result the two strands separate from each other. Each strand acts as template for the synthesis of a new second strand. The synthesis of the new strands is marked by a pair of primers (short sequences of DNA) that are complementary to each of the two DNA strands. The primers are allowed to anneal with the target regions at around 60oC. In the presence of heat stable Taq DNA polymerase, deoxy-nucleotides (dNTPs), $MgCl_2$ and reaction buffer, the synthesis of new strand takes place at 72°C which is the optimum temperature of Taq polymerase. During synthesis or extension step the dNTPs are added to the complementary bases of the target template. At the end of the extension step two molecules of double stranded DNA are formed from a single starting molecule (Fig.3.1). If a similar cycle of heating and cooling is repeated again the newly synthesized double stranded DNA molecules act as templates for further DNA synthesis and from two molecules four molecules of DNA are formed.

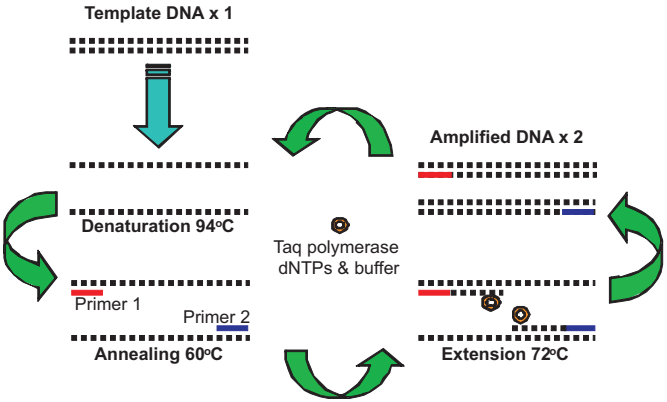


Fig: 3.1. Basic principle of polymerase chain reaction. At the end of one cycle of sequential heating and cooling one molecule of template DNA is amplified to produce two molecules of DNA. In each successive cycle the newly formed (amplified DNA) molecules also act as templates for further amplification.

Sequential repetition of the denaturation, annealing and extension cycles leads to doubling of the DNA molecules with each cycle. In thirty cycles one molecule of DNA can be amplified to produce approximately one billion molecules of DNA. The number of DNA molecules produced from one molecule of target DNA can be calculated by the formula 2^n, where n is the number of cycles.

Size of the new double stranded DNA molecules is defined by the two primers used for amplification. In the initial few cycles the newly synthesized DNA strands are a little longer than the actual boundaries defined by the two primers. However, in the subsequent cycles when more and more of the newly synthesized strands

are used as templates the amplified DNA fragments of exactly the same size, defined by the two primers, are produced.

The initial outburst of amplification in a PCR is gradually reduced to a platue. This is due to the consumption of dNTPs and primers and progressive denaturation of Taq polymerase at very high temperatures. Many other factors like the quality and the quantity of the template DNA, presence of inhibitors of PCR, and the quality of the consumables also play important role in the overall performance of a PCR.

PCR Primers

Primer design

The amplification of DNA essentially requires a short sequence of DNA or "primer" to anneal with the template. Primer or oligo is the main determinant of specificity in a PCR It also has important bearing on the sensitivity of the reaction. Therefore it is essential to understand the basic parameters of designing a PCR primer. The primers may be designed manually or by many computer software packages available for this purpose. The latter also provide simulation data about the performance of the primers.

Primer length

The length of a primer is measured in bases. The usual length ranges from 20-25 bases. The number of bases in a primer is often called "mer" e.g. 20 or 30 mer. A short primer is more likely to anneal non-specifically at sequences that are even partially identical. Increasing the length of the primer increases its specificity but reduces the sensitivity.

Primer ends (5' and 3')

Primers, like any other DNA sequence, have a 5' (5 prime) and a 3' (3 prime) end. For example the primer 5'-CAATGTATCATGCCTCTTTGCA-3' begins at C (5' end) and ends at A (3' end).

The 3' end of the primer is the most critical region as the extension (addition of dNTPs) occurs at this end. The last few bases, especially the last base, at the 3' end are most critical in annealing of a primer. A mismatch at the last base could seriously impair the ability of a primer to anneal. It is also important that the last five bases at the 3' end should not have more than three G or C. The 5' end of the primer is less significant. This end is often used for tagging the primers with compounds like fluorescent dyes etc.

Primer direction (forward and reverse primers)

The two strands of DNA run in opposite directions. Similarly the primer and the template DNA also run in opposite directions. In amplification of double stranded DNA a primer each is used for the two strands i.e. one for the forward strand and the other for the reverse strand. It is customary to name the two primers by their

direction. The primer that anneals with the reverse strand of the DNA is called forward or sense primer while the primer that anneals with the forward strand is called the reverse or ante-sense primer.

DNA is a double stranded molecule but for writing its sequence only the forward strand is written. The primer shown on the forward strand is meant to anneal with the reverse strand. The primer for the other (forward) strand can be understood if the reverse strand is also written. However, to mark the location of the two primers these are often shown on the same strand (Fig 3.3).

Primer sequence

The sequence of nucleotides (G, A, C & T) in a primer that is complementary to the target is called its sequence. In a good primer the numbers of G, A, C and T should ideally be equal. But in practice it is not always possible. However, the total GC content in a primer should remain between 40-60%. Too many G and C can lead to formation of secondary structures and impair the performance. The presence of repeats like ATATAT can lead to non specific annealing.

Primer Tm and annealing temperature

The melting temperature (Tm) is the temperature at which half of the primer is annealed to its target. Tm is dependent on the nucleotide sequence of the primer. Greater the numbers of G and C in a primer higher is its Tm. There are many formulae that can be used for calculation of Tm. For example:

$$Tm = 2 \times AT + 4 \times GC$$

(AT is the sum of A and T nucleotides, and GC is the sum of G and C nucleotides).
The primer designing software and primer vendors also provide data on the Tm of a primer. It is essential that two or more primers being used together in a PCR should have nearly equal Tms. The optimum annealing temperature of a primer is generally 3-5°C below the theoretically calculated Tm.

Primer dimers and secondary structures

Partial similarities in the sequence of primers can make them anneal with each other to form primer dimers. This may include self-dimers (between themselves) or cross-dimers (with other primers) (Fig 3.2). The primer dimers may become difficult to control in a multiplex PCR (PCR with more than two primers). Primers may also form secondary structures like loops and hair pins (Chapter 1). Strong primer dimers and secondary structures may reduce the bioavailability of a primer for the actual target leading to impaired amplification. The problem can become serious when primers are designed manually. The primer designing softwares are useful in tackling this issue.

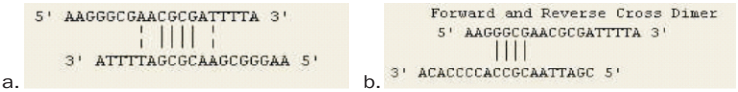


Fig. 3.2. Primer self-dimers (a) and cross-dimers (b).

Primer cross-homology

A good primer should be specific only for its own target. Occasionally a primer may have partial or rarely a complete sequence homology with another site in the genome. This would cause nonspecific amplification. The problem may be overcome by BLAST searching of the newly designed primer sequence (Chapter 1).

Primer designing by computer software

The primers can be designed on computer software or by manual method. For some applications like allele specific PCR the primers are better designed manually. Many commercially available software packages are available with numerous options. A simple free of cost primer designing software "Primer3" is available at the following web address:
http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi

The sequence for which the primers are to be designed can be uploaded as a file or it can be pasted in the software window. Particular regions of the sequence can be marked for inclusion or exclusion or as a target. The software provides numerous options for a large number of primer parameters that can be chosen as per requirement. A simple interface for the Primer3 software called "Primer3Plus" is also available at the following web address:
http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi

Primer labeling

Primers are often required to be labeled with dyes like fluorochromes etc. All such labels are attached at the 5' end which is also the non performing end of the primer. A typical fluorescent dye "FAM" labeled primer would be written as:

5' FAM-CAGTAGCATCTGACTTTGAGCCTCAGGGTCT-3'

Primer purification

During the process of primer synthesis the yield is often contaminated with many short length sequences. This is especially common when primers longer than 30 bases are ordered. The presence of short length sequences can cause non specific amplification. This can consume other ingredients like dNTPs and Taq and may lower the specific amplification. The short length primers can also falsely elevate the final concentration of the working primer solution. In applications where it is critically important to get only specific amplification the primers may have to be purified before use. The primer manufacturers also provide option of gel or HPLC purification of a primer at an extra cost. However, primers shorter than 30 bases usually do not require extra purification.

Primer concentration and dilution

An optimum concentration of primer is critical in a PCR. Too much primer causes it to anneal non-specifically whereas too little primer reduces its availability for

annealing and the amplification. Most PCRs are done with at least one pair of primers. It is important that the two primers of a pair are used in equi-molar ratios. The primer concentration is usually determined by the optical density (OD) at 260 nm. In general 1µg of a primer with 25 bases has an OD of 25. The primer manufacturers provide concentration of primers in µg and OD when reconstituted in one ml of water. The primer is usually supplied in a lyophilized form that is reconstituted in water and is diluted as per requirement. In the lyophilized form it is very stable at room temperature. The lyophilized primer is usually diluted in 1ml of water and it may be stored in aliquots at -20°C for several months or -80°C for several years.

It is convenient to use primer concentration in micro-moles (µM) or pico-moles per micro-liter (pM/µl). The optimum concentration of a primer in most PCRs ranges between 5-10 pM/µl. The following example can explain the dilution of a lyophilized stock primer to make a final concentration of 5pM/µl:

1. OD of the lyophilized stock primer: 13.5 (provided with the stock primer vial).
2. Reconstitute the lyophilized stock primer by adding 1ml water (preferably DNase free).
3. Gently mix by inverting several times.
4. Aliquot and store the stock at -20 to -80°C.
5. To make 100µL of 5pM/µl primer use the following volumes:
   - Volume of the stock primer = 1/OD x volume required

     1/13.5 x 100 = 7.4µl
   - Volume of distilled water to be added: 100-7.4 = 92.6µl
6. Using 1µl of the diluted primer will give a final concentration of 5pM per reaction.
7. To make 10pM/µl concentration, double the volume of the stock primer (14.8µl stock primer + 85.2µl water).
8. Two or more primers to be used in a PCR reaction may also be diluted together in one vial. For example four primers with ODs of 10.5, 13.4, 12.5 and 15.1 may be diluted together in one vial, at concentration of 5pM each in 100µL final volume, as follows:
   - Calculate the volume of the individual primers as above.
   - Add the volume of the four stock primers (1/10.5 x 100) + (1/13.4 x 100) + (1/12.5 x 100) + (1/15.1 x 100).
   - Add distilled water equal to 100 – volume of four stock primers.

Ordering the primer for synthesis

The designed primers are synthesized on a DNA synthesizer. One may synthesize the primers in the lab if a DNA synthesizer is available. But the process is not cost

effective unless very large numbers are required on a regular basis. The best option is to order the primer synthesis from a good commercial company. The request for primer synthesis should include the sequence of the primer, the choice of purification and the synthesis scale. The primers are usually synthesized at a scale of 50 or 200nM. The labeling of the primer, if required, is also indicated.

Taq DNA polymerase

Synthesis of DNA requires a polymerase. Since the temperature in a PCR is raised to 94°C an ordinary polymerase would be denatured at this temperature. The problem is overcome by using bacterial polymerase derived from Thermus aquaticus, that grow in hot water springs. The optimum temperature of Taq polymerase is 72°C.

The Taq polymerase also has 5' to 3' exonuclease activity that can remove any dye attached to the 3' end of a primer. This function is used in removing a dye/quencher attached to the 3' end of a TaqMan® probe used in real time PCR.

Tth DNA polymerase

It is another heat stable polymerase that is derived from Thermus thermophilus. The enzyme has dual function of reverse transcriptase and DNA polymerase. This is useful in cDNA synthesis from RNA and its amplification in a single tube reaction.

High fidelity Taq polymerase

The Taq polymerase lacks 3'to 5' exonuclease activity and therefore is unable to proof read any errors in DNA synthesis. The enzyme Pfu DNA polymerase by virtue of its 3' to 5' exonuclease activity can do the proof reading. It is especially useful in doing PCR of very long templates (high fidelity amplification).

Deoxyneucleotides (dNTPs)

Deoxynucleotides (dNTPs) are the building blocks in new DNA synthesis. These are available from many commercial sources as a set of four nucleotides i.e. dATP, dTTP, dGTP and dCTP. These are usually available in 10 or 100mM concentrations. The four nucleotides are also available as a premixed solution. If a set of four separate nucleotides is purchased then these may be mixed in equal proportions and stored as aliquots at -20 or -80°C.

Magnesium chloride

Magnesium ions increase the solubility of dNTPs and their incorporation in the template. These ions also stimulate Taq polymerase. Therefore the concentration of $MgCl_2$ in a PCR is critical. The amplification can be enhanced by increasing the concentration of Mg. The optimum results are achieved at concentration between 1.5-2.0mM $MgCl_2$.

PCR mix (buffer)

PCR buffer without dNTPs is mostly supplied with Taq polymerase. It may also be prepared in the lab. The following buffer formulation gives consistently good results in a wide range of PCR applications:

Stock Solutions:

2M KCl:	14.9 g/dl
1M Tris:	12.1 g/dl
1M $MgCl_2$:	9.52 g/dl
1M Spermidine:	14.52 g/dl

10 X PCR Buffer

2M KCl:	1.25 ml
1M Tris (pH 8.3):	0.5 ml
1M $MgCl_2$:	75 µl
Gelatin (300 bloom):	5 mg
Distilled water:	3.2 ml

PCR mix (1 ml)

10 X PCR buffer:	50 µl
10mM dNTPs (all):	25 µl
1M Spermidine:	1 µl
Distilled water:	925 µl

Inhibitors and enhancers of PCR

Haemoglobin, heparin, porphyrin, SDS, phenol, and proteinase-K are potent inhibitors of PCR. SDS, and phenol must be completely removed during DNA extraction. Proteinase K can be inactivated by heating. PCR may also be inhibited by chocolate that might contaminate the DNA extracted from a mouthwash. Enhancers of PCR include formamide, DMSO, tetramehylammonium chloride (TMAC), polyethylene glycol (PEG), glycerol, tween 20 and 7 deaza-dGTP.

Setting up a PCR

There is a wide variety of PCR applications that differ in the reaction conditions. An example of amplification of β-globin gene containing C-T polymorphism is described. The sequence of the gene was downloaded (GenBank NG_000007.3) and the primers flanking the C-T polymorphism were designed using the Primer3 software (Fig 3.3).

TTTAGACATAATTTATTA**GCATGCATGAGCAAATTAAGA**AAAACAACAACAAATGAA
TGCATATATATGTATATGTATGTGTGTA**[C/T]**ATATACACACATATATATATATAT
TTTTTCTTTTCTTACCAGAAGGTTTTAATCCAAATAAGGAGAAGATATGCTTAGAAC
CGAGGTA**GAGTTTTCATCCATTCTGTCCTG**TAAGTATTTTGCATATTCTGGAGACGC

Fig. 3.3. PCR amplification of β-globin gene (GenBank NG_000007.3) containing C-T polymorphism. The primers flanking the polymorphism were designed using the Primer3 software

Locus:	β-globin gene
GenBank accession:	NG_000007.3
Forward primer:	5'-GCATGCATGAGCAAATTAAGA
	(Length: 21, Tm: 59°C, GC 38%)
Reverse primer:	5'-CAGGACAGAATGGATGAAAACTC
	(Length: 23, Tm: 60°C, GC 43%)
Amplified product size:	179bp

PCR master mix

For 10 reactions (7 test samples + 1 positive control + 1 negative control + 1 reagent blank) make 260µl master mix (25µl/reaction + 10µl extra for pipetting errors). Mix the following in a 0.5ml plastic tube:

- PCR buffer with dNTPs: 240µL (24µL per reaction)
- Forward primer 10µL (5pmol/µL) (1µL per reaction)
- Reverse primer 10µL (5 pmol/µl) (1µL per reaction)
- Taq polymerase (5 units/µL) 1.0µL (0.1µL/reaction)
- Template DNA (~200ng/µl): 1µL/reaction

1. Take 10 x 0.2ml PCR reaction tubes in a rack and label these 1-10.
2. Dispense 25µL master mix to each of the ten labeled tubes.
3. Add 1µL test DNA samples to tubes 1-7.
4. Add 1µL control DNA positive for C-T polymorphism in tube 8.
5. Add 1µL control DNA Negative for C-T polymorphism in tube 9.
6. Add 1µL water instead of DNA in tube 10.
7. Firmly close the lids of the tubes and put these in thermal cycler.
8. Programme the thermal cycler and run PCR as follows:
   a. Initial denaturation 94°C for 5 minutes.
   b. Thirty cycles each of denaturation at 94°C for 1 minute, annealing at 60°C for 1 minute and extension at 72°C for 1 minute.
   c. Final extension at 72°C for 3 minutes.
9. Electrophoresis:
   - 10 x 10 cm 6% polyacrylamide, 30 minutes at 150 volts.
10. Staining:
    - 0.1% silver nitrate

DNA Thermal Cycler

Thermal cycler is a machine that can be programmed to heat and cool rapidly for PCR amplification. A wide variety of machines are available from different sources. The choice of machine depends on the requirement of a lab. An average thermal cycler can heat or cool at about 1°C per second. More expensive machines with very fast heating or cooling (ramping) are also available. Very fast ramping can shorten the time required in thermal cycling. The fast machines also require special thin walled PCR tubes that allow heat transfer at an equally fast rate. Most PCR machines have 96 well capacities. Smaller machines with capacity of 32 or 48 well are also available.

Hot start PCR

At the start of PCR when the temperature in the reaction vial is rising the primers may anneal non specifically with a partially denatured template resulting in nonspecific amplification. This may be prevented if the Taq polymerase is added after the template is completely denatured. In a hot start PCR Taq polymerse is added when the contents of the PCR reaction mix have reached above 80°C. The process may be helped by using AmpliwaxTM. A small wax bead is placed above the reaction mix in a PCR tube. Taq polymerase mixed in buffer is added on top of the wax bead. When the temperature in the vial reaches above 80°C the wax melts and the Taq polymerase is mixed with the reaction mixture.

Nested PCR

When the target DNA is low in concentration the amplification may be done in two steps. In the first step a pair of flanking primers is used to amplify the target. In the second step the amplified product of the first step is re-amplified with another set of primers located within the outer/flanking primers used in the first step. The nested PCR is extremely sensitive. The first step amplified product may easily contaminate the other PCR reagents and may be very difficult to handle.

Bibliography

1. Newton CR and Graham A (1994) PCR. Bios Scientific Publishers, Oxford, UK.
2. Saiki RK, Gelfand DH, Stofel S, Scharf R, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487
3. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM (2007) Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 35: W71-W74.

4. Synthesis of cDNA

RNA cannot be amplified like DNA. However, it can be converted to copy DNA or cDNA which can be amplified like any other DNA sequence. The process requires "primer" complementary to the target RNA sequence, enzyme called reverse transcriptase, deoxy-nucleotides (dNTPs), magnesium ions and reaction buffer. The end result is formation of cDNA strand that is complementary to the RNA template sequence. The process is also called reverse transcription.

The primer

cDNA may be synthesized by any of the three primers:

1. Gene specific primer
2. Random hexamers
3. Oligo dT primer

Gene specific primers are complementary to the target itself. These can be used in all types of applications but are especially suitable for cDNA synthesis and PCR amplification in the same tube. These are also useful for synthesis of long stretches of cDNA.

Random hexamers are a mixture of all possible combinations of six nucleotide pieces of DNA. These are useful when only small stretches of cDNA or multiple target cDNAs are to be synthesized (Fig 4.1).

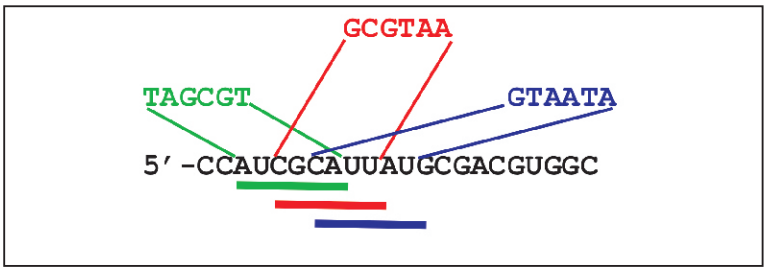


Fig. 4.1 Random hexamers are mixture of all possible combinations of six nucleotides.These anneal randomly to their target RNA and produce relatively short stretches of cDNA.

Oligo dT primers are designed to anneal with the poly-A tail attached at the 3' end of mRNA molecules. Since mRNA constitutes only 1-2% of the total RNA, the yield of cDNA from oligo dT is low. These are useful for synthesizing long stretches of cDNA or when multiple targets are to be synthesized.

Reverse transcriptase

Reverse transcriptase is an enzyme that can be used for synthesizing DNA copy of RNA. The most commonly used reverse transcriptase is derived from Moloney murine leukaemia virus (MMLV). It is an RNA dependent DNA polymerse that can synthesize DNA complementary to the RNA sequence. In the presence of a primer

the DNA synthesis starts from 5' to 3' direction. A thermostable enzyme derived from Thermus thermophilus (Tth) has dual activity of reverse trancriptase and DNA polymerase. It can be used for reverse transcription and subsequent PCR in the same reaction tube.

Example

An example of cDNA synthesis from the Human Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) mRNA (GenBank NM_002046) is described. The primers were designed by Primer3 plus software (Fig. 4.2).

TGCGTCGCCAGCCGAGCCACATCGCTCAGACACCATGGGGAAGGTGAAGGTCG **GAGT**
CAACGGATTTGGTCGTATTGGGCGCCTGGTCACCAGGGCTGCTTTTAACTCTGGTAA
AGTGGATATTGTTGCCATCAATGACCCCTTCATTGACCTCAACTACATGGTTTACAT
GTTCCAATATGATTCCACCCATGGCAAATTCCATGGCACCGTCAAGGCTGAGAACGG
GAAGCTTGTCATCAATGGAAATCCCATCACCATCTTCCAGGAGC **GAGATCCCTCCAA**
AATCAAGTGGGGCGATGCTGGCGCTGAGTACGTCGTGGAGTCCACTGGCGTCTTCAC

Fig. 4.2. PCR primers for the human Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH), mRNA (GenBank NM_002046.3). The reverse primer (blue) was used as gene specific primer for reverse transcription.

cDNA synthesis

RT Primer (gene specific): 5'-TTGATTTTGGAGGGATCTCG

Reaction conditions:

- 5 X reaction buffer: 4 µl
- 10 mmol dNTPs mix: 2 µl
- RT Primer (10 pmol/µl): 1 µl
- MMLV- Reverse Transcriptase (200 U/µl): 1 µl
- RNase inhibitor (20 U/µl) 1 µl
- Deionized water: 13 µl
- RNA: 2 µl
- Incubation: 42°C for 60 minutes
- RT inactivation: 70°C for 5 minutes

In order to reduce cost the typical reaction volumes may be reduced to suit individual requirements.

PCR amplification

Forward primer: 5'-GAGTCAACGGATTTGGTCGT
(Length: 20, Tm: 60oC, GC 50%)
Reverse primer: 5'-TTGATTTTGGAGGGATCTCG
(Length: 20, Tm: 60oC, GC 45%)
Amplified product size: 238bp

PCR master mix

- PCR buffer with dNTPs: 20 µl
- Forward primer 1 µl (5pmol/µl)

- Reverse primer 1 µl (5 pmol/µl)
- Taq polymerase (5 U/µl) 0.1 µl
- cDNA 3 µl

Thermal cycling:

- Initial denaturation 95°C for 5 minutes.
- Thirty cycles of:
  - Denaturation 95°C for 1 minute
  - Annealing 58°C for 1 minute
  - Extension 72°C for 1 minute 30 seconds
- Final extension 72°C for 3 minutes

Electrophoresis: 6% polyacrylamide at 150 volts for 40 minutes
Staining: 0.1% silver nitrate

Bibliography

1. Lowe K (2001) First strand cDNA synthesis. Oxford Practical Series. Oxfor University Press.
2. Green MR, Sambrook J (2012) Molecular Cloning: A Laboratory Manual. *(Fourth Edition). Cold Spring Harbor Laboratory Press.*

5. Gel Electrophoresis of DNA

Electrophoresis is used for the separation of DNA, RNA, or other charged molecules like proteins by an electric current applied to a gel matrix. In PCR applications it is usually done for separation and visualization of DNA after PCR or sequencing. The electrophoresis gel is a porous matrix that allows movement of DNA through its pores. DNA fragments with large differences in size are typically separated on agarose gels whereas small differences in DNA fragment sizes are resolved on polyacrylamide gels.

The nucleic acids have a negative charge due to the sugar phosphate backbone of the molecule. This allows movement from the negative to the positive electrode. The double-stranded DNA fragments naturally behave as long rods, so their migration through the gel corresponds to their size. Single-stranded DNA or RNA tends to form secondary structures with complex shapes and migrate through the gel in a complicated manner. This can be overcome by adding denaturing agents like formamide or urea to remove secondary structures and cause them to behave as long rods.

Agarose gel electrophoresis

Agarose is composed of long un-branched chains of carbohydrate. Since there are no cross links between chains these gels have large pores and are suitable for separation of larger molecules of DNA. NuSieve agarose is a special product with smaller pore size and better resolution. The agarose gels are typically run in horizontal tanks after submerging in a buffer solution hence the name submarine gels (Fig. 5.1). The agarose gels are unable to withstand high voltage because the heat generated may cause melting of the gel.

Preparing 2% agarose gel (14 x 10 x 0.5 cm)

1. In a 500 ml beaker take:
   - Agarose 1.6 g
   - 1 x AGB buffer 80 ml
2. Cover with a cling film.
3. Heat in a microwave oven for 2 minutes at high power.
4. Seal the ends of a gel tray with scotch-tape.
5. Pour the molten agar on the gel tray.
6. Place the appropriate sample comb in the gel tray.
7. Leave to cool for 30-60 minutes. The gel will solidify and appear opaque grayish in colour.
8. Make sure to remove the scotch tape before running the gel.

To prepare 4% NuSieve agarose gel (14 x 10 x 0.5 cm) take the following in a 500 ml beaker:

a. Agarose: 1.6 g
b. NuSieve agarose: 1.6 g
c. 1 X AGB buffer 80 ml

Running the agarose gel

1. Fill the electrophoresis tank with 1 x TBE buffer.
2. Place the gel in the electrophoresis tank while it is still in the casting tray.
3. Carefully remove the comb(s) from the gel. If the comb is removed hurriedly or the gel has not solidified completely the well floor may be damaged.
2. Prepare the amplified DNA for loading by mixing 20µl amplified product and 5µl loading dye. The amplified DNA must be handled with care as it is a potent source of contamination for other PCR reactions. It is advisable to use a separate pipette marked as "amplified DNA only".
3. Carefully load 20-25µl of the DNA and loading dye mixture in to the wells. The DNA in the sample quickly sinks to the floor of the well because of the glycerol in the loading dye.
4. Start electrophoresis at ~150 volts for 30-60 minutes depending on the size of the DNA fragments to be resolved. Use of higher voltage can cause melting of the gel.
5. Progress of electrophoresis can be monitored by movement of the blue coloured loading dye. Bromophenol blues moves in the agarose gel at approximately the speed of a 200bp DNA fragment.

Staining of agarose gels

The agarose gels are stained with ethidium bromide which is an intercalating dye that binds to double stranded DNA. It fluoresces when exposed to ultraviolet light. The reddish orange fluorescence of ethidium bromide increases 20 fold after binding to double stranded DNA. Ethidium bromide is a carcinogen and mutagen therefore it should be handled after wearing gloves.

1. Add 3-5 drops of 1% ethidium bromide solution to 500 ml of water in a staining tray.
2. Carefully remove the gel from the tray and submerge it in the stain.
3. Keep in the stain for 30 minutes.
4. Transfer the gel to 500 ml clean water and keep for 5-10 minutes. This will wash the excess background satin from the gel.
5. Place the gel on UV transilluminator. The amplified DNA fragments are seen as reddish orange bands (Fig. 5.2).

A permanent record of the gel is kept by taking a picture with a camera or gel documentation system.

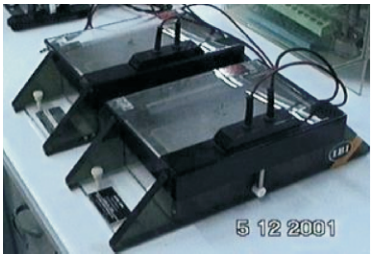


Fig. 5.1. Submarine gel electrophoresis tank.

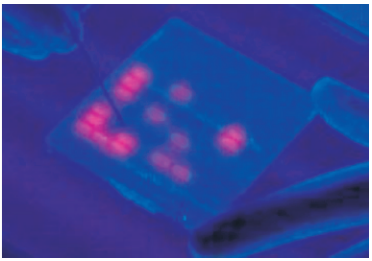


Fig. 5.2 Ethidium bromide stained agarose gel. The amplified DNA appears as reddish orange bands.

AGB Buffer (10 X)

Tris:	48.4 g
$NaAC.3H_2O$:	27.2 g
Disodium EDTA:	0.744 g
Glacial acetic acid:	12 ml
Deionized water:	up to 1L

TBE Buffer (10 X)

Tris:	108 g
Boric acid:	55 g
Na_2EDTA:	9.3 g
Deionized water:	up to 1L

Loading dye for agarose gels

• Bromophenol blue	5 mg
• Glycerol	10 ml
• Deionized water	up to 100 ml

Ethidium bromide (1% stock solution)

• Ethidium bromide	1 g
• Distilled water	100 ml

Acrylamide gel electrophoresis

Acrylamide is a white crystalline powder that forms polymers after dissolving in water. The acrylamide polymers can be cross-linked by addition of bisacrylamide. As a result of cross-linking the pore size of the gels also decreases. This makes acrylamide gels most suitable for separation and resolution of DNA molecules that may differ in size by only one base pair. The acrylamide gels can withstand high temperatures and hence are suitable for applications requiring high voltage. Acrylamide gels are typically run in a vertical position using discontinuous buffer compartments (Fig. 5.3).

Acrylamide in solution becomes viscous but does not form a gel. The cross-linked acrylamide polymers can be solidified to form a gel by adding ammonium persulphate (APS) and N,N,N',N'-tetramethylethylenediamine (TEMED) in equimolar ratio. The concentration of APS and TMED determine the rate of gel formation and its turbidity and elasticity.

Acrylamide should be stored in a cool dark and dry place to reduce autopolymerisation and hydrolysis. Acrylamide is neurotoxic that is absorbed through skin. Its effect is cumulative and toxicity may develop over prolonged exposure. Gloves must be worn when using acrylamide.

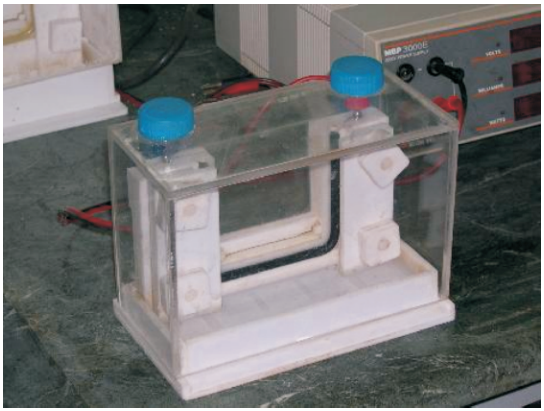


Fig. 5.3. Vertical polyacrylamide gel electrophoresis tank.

Preparation of 6% polyacrylamide mini gel (10 x 10 x 0.1 cm)

1. Make 6% acrylamide solution by mixing the following:
   - Acrylamide 28.5 g
   - N, N' methyl bis-acrylamide 1.5 g
   - TBE buffer (10 X) 50 ml
   - Distilled water up to 500 ml
2. Wash 10 x 10 cm glass plates with soap and water.
3. Dry the plates completely with cotton gauze.
4. Put 1-2 ml ethanol on the surface of the plates facing the gel and wipe the surfaces thoroughly with dry piece of gauze to remove any grease etc.
5. Assemble the glass plates as per instructions of the manufacturer.
6. Place the required comb between the plates.
7. Take 10 ml 6% acrylamide in a small beaker.
8. Add 100µl 10% ammonium persulphate (APS) not older than one week.
9. Add 20µl TEMED.
10. Gently mix and fill the polymer in a 10 ml syringe.
11. Attach a 21 gauge butterfly needle set and gently push the plunger to fill the polymer in the tubing. Remove bubbles in the tube if necessary.

12. Gently pour the acrylamide between the plates before the gel polymerization starts (within 2-3 minutes).
13. Remove bubbles if any by gently tapping the plates.
14. Allow up to 30 minutes for complete polymerization to occur.
15. For larger polyacrylamide gels increase the polymer and the catalysts accordingly. For example 16 x 20 x 0.1 cm gel would take 32 ml polymer. Make about 35 ml polymer in a beaker and add 350µl APS and 70µl TMED.
16. The larger gel plates must be cleaned thoroughly to leave no trace of grease or dirt on the plates. When preparing very large gels one of the plates (not both) may be coated with silicon to allow easier separation at later stage.

Running acrylamide gel

1. Remove the gel plates from the casting assembly and fix it on the gel tank as recommended by the manufacturer.
2. Fill the upper and the lower buffer compartments with 1x AGB buffer.
3. Gently pull the comb out and wash the wells with buffer to remove any residual un-polymerized acrylamide. A syringe with 21 gauge butterfly needle may be used for washing wells.
4. Prepare the samples for loading by mixing 2-4µl amplified product (depending on the quality of amplification) and 2-3µl of loading dye (0.05% xylene cyanol and bromophenol blue in formamide). The samples may be prepared in 96 well ELISA plates or 0.2ml plastic tubes.
5. Load 3-4µl of the prepared sample with a long nose plastic tip. The amplified DNA must be handled with care as it is a potent source of contamination for other PCR reactions. Use separate pipette marked as "amplified DNA only".
6. Run the gel at 150-200 volts for 15-45 minutes depending on the fragment size of the amplified products.

Staining of acrylamide gels

Silver staining

1. Remove the gel from casting assembly.
2. Carefully separate the plates taking care to avoid tearing of the gel.
3. Mark the first well or the side of the gel by cutting a small piece from the left lower corner of the gel.
4. Put the gel in 0.1% silver nitrate solution for 15-20 minutes. The gel may be kept in stain while it is still on the glass plate.
5. Discard the stain and wash the gel in plenty of tap water.
6. The stain may be reused if kept in dark brown bottles. However it looses potency after exposure to bright light.
7. Prepare fresh developing solution by adding 75µl formaldehyde to approximately 100 ml 1.5% NaOH. Stock NaOH is stable at room

temperature but once formaldehyde is added it must be used within one hour.

8. Submerge the gel completely in developing solution.
9. In approximately 5-10 minutes the bands of amplified DNA can be seen on the gel. The background of the gel also becomes light yellowish brown. The background colour could become very dark if the gel is kept for too long in the developer solution. The exact developing time can be learnt by trial and error.
10. Discard the developing solution and wash the gel in plenty of water when the DNA bands are clearly seen.
11. Cut a piece of filter paper slightly larger than the gel itself and lay it flat on the gel surface. Gently pick the filter paper along with the gel that sticks to its surface.
12. Place the gel and the filter paper on a gel dryer making sure that the gel faces towards the front.
13. Dry the gel under vacuum for 20-30 minutes at 80°C.
14. The dried gel can be pasted in a record book for long term storage after trimming its margins.
15. The gel may be photographed if a gel dryer is not available.

Ethidium bromide staining

The acrylamide gels can also be stained in ethidium bromide as described in section of staining agarose gels.

Denaturing Gradient Gel Electrophoresis (DGGE)

The single stranded DNA tends to form secondary structures (Chapter 1). The movement of single stranded DNA in acrylamide gels, in addition to the size of the fragments, is also dependent on its base composition. The base composition determines the amount and the shape of the secondary structures.

When a DNA fragment 200-700bp in length is run on acrylamide gel with increasing concentration (gradient) of a denaturant like urea it would initially move according to the molecular weight. As it moves into the higher concentration of the denaturant it reaches a point where the DNA starts to melt (the two strands begin to separate). The resulting single strands of DNA also start to develop secondary structures. The partial melting and formation of secondary structures severely retard the progress of the molecule in the gel. A single base pair change in the DNA can cause significant mobility shift. In this way different alleles can be identified by the differences in mobility on a gel with a gradually increasing denaturant concentration.

Constant temperature of 60°C, formamide (0-40%) and gradually increasing concentration of urea from 0-7 M are usually used as denaturants in DGGE.
Preparation of denaturing gradient gel
Prepare the following stock solutions:

1. Acrylamide (40%) and bis-acrylamide (1.07%)
2. 20 x TAE gel running buffer (pH 7.4) containing 800 mM Tris base, 400

mM sodium acetate, and 20 mM EDTA.

3. 80% denaturant stock solution containing 7% acrylamide, 32% formamide, 5.6 M Urea and 1 x TAE buffer.
4. 0% denaturant stock solution containing 7% acrylamide, and 1 x TAE buffer.

Formamide is deionized before use by gently stirring with Dowex AG50W (20-40) mesh mixed bed resin. The stock solutions of 0% and 80% denaturent are used to prepare the gradients of varying strengths. Gels measuring 20 x 16 cm and 1.5 mm thick are poured by a gradient mixer. Ammonium persulphate (100µl of 10% stock solution) and 10µl of TEMED are used as gel polymerization catalysts.

To make a 25-50% gradient of the denaturant the required volumes of 0% and 80% stock solutions are shown in Table 5.1.

Denaturant	25% (lower limit)	50% (upper limit)
0% denaturant	13.8 ml	7.5 ml
80% denaturant	6.2 ml	12.5 ml
Total volume	20 ml	20 ml

Table. 5.1. The volumes of 0% and 80% denaturant solutions to be used for preparation of denaturing gradient gels.

Making gradient gel

The gradient mixer is made of two chambers connected through a tube close to the bottom. The connection is normally closed. One of the chambers (downstream) has an outlet where tubing for the out flowing polymer is attached. This outlet is also kept closed. The required amount of high denaturant mix (e.g. 50%) is added to the out flowing chamber and the low denaturant mix (e.g. 25%) is added to the other chamber. It is important that the volumes of the two denaturant mixtures should be equal. A small magnetic stirrer is placed in the out flowing chamber. The gradient mixer is placed at a height of about 12 inches above the gel casting assembly. A 23 gauge butterfly needle is attached at the outlet.

Before pouring the gel the tap between the two chambers is opened and the magnetic stirrer is started Ammonium persulphate (100µl of 10% stock solution) and 10µl of TEMED are added to the polymer in each of the chambers. The tap at the outlet orifice is opened and bubbles in the needle are removed. The butterfly needle is placed between the two glass plates and the polymer is allowed to fill the space between the plates. As the polymer flows from the upstream chamber to the out flowing chamber the polymer in the two chambers is gradually mixed and this forms a gradual gradient with highest concentration at the bottom of the gel and the lowest concentration at the top of the gel.

Running conditions for DGGE

Electrophoresis is carried out on a vertical polyacrylamide gel running system. A specially designed acrylic chamber to hold the gel assembly, heater and buffer circulation device is used (Fig 5.4). The samples are loaded on the gel in the usual way. In each well up to 12µl of the sample and loading dye (0.05% xylene cyanol in formamide) can be used. Electrophoresis is carried out at 50V for 16 hours. Throughout the procedure buffer temperature is maintained at 60°C. At the end of the run the gel is removed and it is stained with silver nitrate or ethidium bromide (Fig 5.5).

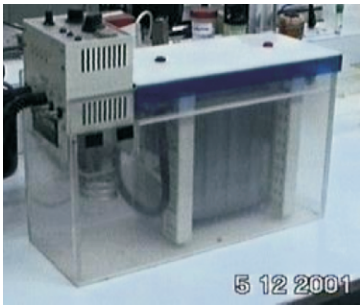


Fig. 5.4. Apparatus for denaturing gradient gel electrophoresis. The main buffer compartment contains an immersion heater and a buffer circulation pump that maintains a constant temperature of 60°C throughout the electrophoresis run.

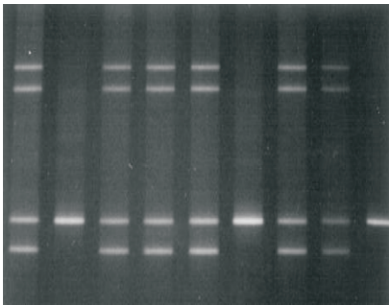


Fig. 5.5. Ethidium bromide stained DGGE.

Single Strand Conformation Polymorphism (SSCP)

Single stranded DNA tends to form secondary structures which are dependent on its nucleotide composition. The secondary structures interfere with the mobility of the strand in a polyacrylamide gel. Two single-stranded DNA fragments with only a single nucleotide difference would migrate at different rates. This forms the basis of SSCP.

Single stranded DNA can be produced by first amplifying DNA and then denaturing it to form single strands. Alternatively single stranded DNA can be produced by asymmetric PCR in which one primers of a pair is used in excess. In the initial cycles of PCR double stranded DNA is produced but as the primer with lower concentration is consumed the primer in excess causes extension of only a single strand. The usual fragment size for SSCP is between 150-300 bases. SSCP is done on non-denaturing polyacrylamide gels that are stained in silver nitrate.

SSCP is a useful technique for screening of SNPs and unknown mutations. SSCP is unable to tell the base composition of DNA unless control DNA fragments of known composition are also run parallel to the unknown samples.

Capillary electrophoresis

Electrophoresis of DNA can be done in very narrow bore long and flexible capillaries. The capillaries are filled with an electrolytic solution and its ends are dipped in buffer chambers. The sample injection and running is done by electric current. The electrophoresis is usually done for short periods at 5000 to 30,000 volts (Fig. 5.6).

Capillary electrophoresis is especially useful for automated fragment analysis and genomic sequencing of fluorescent labeled amplified products. Many automated genetic analyzers are commercially available that allow simultaneous running of several samples. These analyzers are extremely efficient in fragment analysis and genomic sequencing applications.

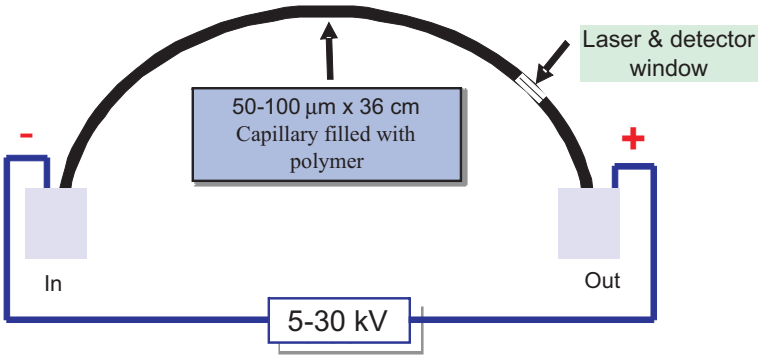


Fig. 5.6. Diagrammatic representation of capillary electrophoresis.

Bibliography

1. Fodde R and Losekoot M (1994) Mutation detection by denaturing gradient gel electrophoresis (DGGE). Hum Mutation 3: 83-94.
2. Grompe M (1993) The rapid detection of unknown mutations in nucleic acids. Nature Genetics 5: 111-117.
3. Cai SP, Kan YW (1990) Identification of the multiple -thalassaemia mutations by Denaturing Gradient Gel Electrophoresis. J Clin Invest 85: 550-553.
4. Constantinos Z (2009, January 18) Capillary Electrophoresis. SciTopics. Retrieved July 2, 2012, from http://www.scitopics.com/Capillary_Electrophoresis.html

6. Real Time PCR

In a PCR the number of DNA molecules double with each cycle. The initial outburst of amplification is reduced to a platue in the last few cycles of the PCR due to the consumption of dNTPs and primers and progressive denaturation of Taq polymerase. The process of amplification can be examined in real time if the amplified DNA can be measured while it is being synthesized.

In a conventional PCR the amplified DNA is examined at the end of PCR whereas in a real time PCR the amplification is monitored (visualized) after each cycle and the result is plotted on a linear or a logarithmic scale. The cycle number at which the amplification first appears is dependent on the quantity of target DNA. The amplification appears earlier when the target DNA is more and it appears late with decreasing concentration of target DNA molecules. This makes real time PCR an ideal tool for quantitative estimation of DNA.

Another advantage of real time PCR is that there is no need for post PCR processing of the amplification products. The amplified product of PCR is a potent source of contamination for other PCR reactions. In a real time PCR since no post amplification processing is required the chances of contamination are also minimized.

Monitoring of amplification

SYBR green method

SYBR green is a fluorescent dye that binds to minor groove of double stranded DNA. It gives little fluorescence when in solution but emits a strong fluorescent signal after binding with double-stranded DNA. In a real time PCR the amplification can be monitored by SYBR green. As the amplified double stranded DNA accumulates in the reaction tube the fluorescence also increases. The main drawback of SYBR green is that it also binds to DNA that is amplified due to non-specific annealing of the primers or due to the formation of primer dimers. SYBR green method is also not suitable for analysis of multiplex PCR products.

The real time PCR machines can also determine the melting point (Tm) of the amplified product by a process called melting curve analysis. Since Tm of double stranded DNA depends on its base composition and size of the DNA fragment, specific and nonspecific products or primer dimers can be differentiated by comparing their Tm values.

Molecular probe methods

The amplification may also be monitored by a more specific method employing fluorescent labeled short sequences of DNA (probe) complementary to the target DNA of interest. Several methods have been developed to effectively use molecular probes for monitoring PCR.

The fluorescence emitted in the molecular probe methods is directly proportional

to the annealing of the probe. Therefore the methods are highly specific unless the target has a cross homology with another sequence in the template DNA.

Fluorescence Resonance Energy Transfer (FRET) Probe method

In the FRET probe method the target is amplified by the usual pair of primers. The FRET probes are a pair of probes that anneal to the inner region of the target in a head to tail configuration. The upstream probe has a fluorescent dye (donor) at the 3' end whereas the downstream probe has another fluorescent dye (acceptor) attached at the 5' end. If the target DNA is present it is amplified by the two primers and the two FRET probes also anneal to the target. The head to tail configuration of the FRET probes ensures that the fluorescent dye at the 3' end of the upstream probe comes in close proximity to the fluorescent dye at the 5' end of the downstream probe. On excitation by the light of appropriate wavelength the energy from the upstream probe (λ1) is transferred to the downstream probe (λ2). The fluorescent dye on the downstream probe after absorbing energy emits this in the form of light of yet another wavelength (λ3). The later is measured by the instrument and is directly proportional to the amount of amplification.

Molecular Beacon method

The target DNA is amplified by the usual pair of primers. The molecular beacon probe is designed to be complementary to the inner region of the target DNA. At each end of the probe 5-7 nucleotides, complementary to each other, are inserted to ensure a hairpin structure (stem) of the probe (Fig. 6.1). At each end of the probe a fluorescent reporter dye and a quencher are also attached. In the non-annealed probe the reporter and the quencher are very close to each other and the fluorescence from the reporter dye is quenched. As a result of the annealing of the probe to its target the hairpin structure opens up and separates the reporter from the quencher. The net result is that on excitation by the light of appropriate wavelength the reporter dye starts emitting fluorescence. The later is directly proportional to the amount of amplification.

Fig. 6.1. Molecular Beacon probe in the un-annealed and the annealed state.

TaqMan® probe method

The TaqMan® probe is a short sequence of DNA complementary to the internal region of the PCR target DNA. The 5' end of the probe is labeled with a fluorescent dye (reporter) while its 3' end is labeled with a fluorescence quencher (suppressor). The quencher may be a high energy fluorescent dye like rhodamine (TAMRA) or a non-fluorescent chemical quencher e.g. DABCYL and black hole

quenchers (BHQ). In the intact probe the reporter dye and the quencher are in close proximity to each other. This suppresses the fluorescence of the reporter dye.

In the TaqMan® method of real time PCR the target DNA is amplified by a pair of primers. The reaction mixture also contains the TaqMan® probe that emits no fluorescence when it is not annealed to its target. If the target DNA is absent there is no amplification and the probe does not emit any fluorescence. When the target DNA is present it would be amplified. During amplification the probe, being complementary to the internal region of the target, also anneals to the target. The Taq polymerase used in PCR has 5' to 3' exonuclease activity which removes the probe from the target in a stepwise manner. First the reporter dye is removed then the probe itself and finally the quencher is also removed (Fig 6.2). As a result the reporter dye is relieved from the quencher and starts emitting fluorescence on excitation by the light of appropriate wavelength. The net increase in fluorescence from the reaction is directly proportional to the amount of probe annealed to the target and hence the amplification.

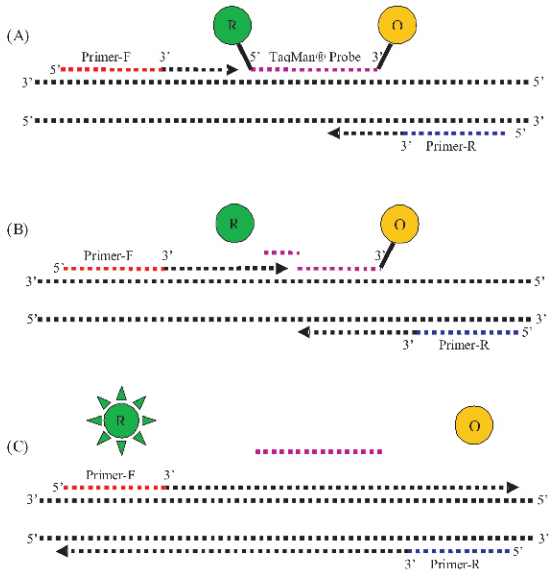


Fig. 6.2. Real time PCR by TaqMan® probe method. (A) The target is amplified by primers F & R. TaqMan® probe is annealed to the target and there is no fluorescence from the reporter (R) because of the close proximity to the quencher (Q). (B) As the primer extension takes place and reaches the 5' end of the probe, the reporter (R) is released due to the 5' to 3' exonuclease activity of Taq polymerase. (C) Continued extension of the new strand removes the probe and finally the quencher (Q). As the distance between the reporter and the quencher increases the reporter dye gives a bright fluorescence on excitation that is measured by the instrument. The quantity of fluorescence is directly proportional to the amount of target and the amplification.

TaqMan® probe designing

The TaqMan® probe is designed while keeping the following parameters in mind:

1. Tm of the probe should be 10°C higher than that of the primers.
2. Runs of identical nucleotides especially "Gs" should be avoided.
3. The G+C content should be 30-80%.
4. There should be more Cs than Gs.
5. There should be no G at the 5' end.

Primers for TaqMan® assay

In addition to the general requirements of primer designing (Chapter 3) the primer pair used with the TaqMan® probe should have the following additional properties:

1. Should have no runs of four or more Gs (any nucleotide).
2. Should have no more than two G+C at the 3' end.
3. Should have no G at the 5' end (A or C is preferred).
4. The amplicon size should range between 50-150bp (max 400).
5. Preferably should span the exon-exon junctions in cDNA.

The TaqMan® probes are best designed by computer software specially made for this purpose. Many such softwares are available commercially. The real time PCR machine vendors may also provide the software with the machine. The sequence of interest is uploaded in the FASTA format or it can be cut and pasted in the software window. The software offers several design options and provides a list of possible primers and the probe combinations. The best combination is chosen for a particular application.

Choice of reporter and quencher

The reporter and the quencher pair for a probe should be compatible with each other chemically as well as for the wavelength of fluorescent light. A wide variety of reporter fluorescent dyes are available to choose from. Table 6.1 gives a list of the reporter dyes with the wavelengths of incident and emerging light. An appropriate dye can be chosen depending on the capability of the machine on which the probe would be used.

The quencher may be fluorescent e.g. TAMRA or chemical e.g. DABSYL and black hole quenchers (BHQ). TAMRA as a quencher has a drawback of giving high background fluorescence. DABCYL also has a limitation of having poor spectral overlap between itself and the reporter. Black hole quenchers (BHQ1, BHQ2 and BHQ3) have been developed to overcome these problems. Table 6.1 also provides a list of compatible quenchers for the relevant reporter dyes.

When choosing probes for a multiplex real time PCR the reporter dyes should be selected so that their excitation wavelengths are clearly separate with minimal overlap.

Passive reference dye

Some real time PCR machines require addition of a passive reference dye in the reaction mixture. The passive reference dye e.g. ROX is included in the reaction master mix. It does not participate in the 5' nuclease reaction but it serves as an internal reference for background fluorescence emission.

Table 6.1. Wavelengths of the incident and the emergent light of some commonly used reporters and the quenchers for TaqMan® probes.

Reporter dye	Incident light wavelength	Emergent light wavelength	Compatible quenchers
FAM/SYBR Green	492	516	TAMRA, DABCYL, BHQ1
TET	517	538	TAMRA, DABCYL, BHQ1
HEX/JOE/VIC	535	555	TAMRA, DABCYL, BHQ1
Cy3	545	568	BHQ2
TAMRA	556	580	BHQ2
ROX/Texas Red	585	610	BHQ2
Cy5	635	665	BHQ3

SYBR green method

SYBR green non-specifically binds to double stranded DNA and is frequently used in real-time PCR reactions. It fluoresces very brightly when it is bound to double stranded DNA. The non-specific binding of SYBR green is a disadvantage in real time PCR as the dye could also bind to the non-specific PCR amplified products and the primer dimers. The problem can be overcome by a special procedure called "melting curve analysis".

Another disadvantage of the SYBR green method is that the larger amplified products give a stronger signal. The SYBR green method is also unsuitable for multiplex PCR unless it is combined with melting curve analysis.

The SYBR green method may be used after extensive optimization in detection of single amplicons with no non-specific amplification and primer dimer formation.

Melting curve analysis

The real time PCR machines using the protocol for SYBR green based methods also provide option for melting curve analysis. At the end of the PCR the temperature of the heating block is raised in steps of 1°C. At each step the Tm of the reaction products is measured by the amount of fluorescence emitted. At the end of the analysis the software plots the Tm readings against the temperature. The presence of a single or more than one amplified products can be identified by their respective Tms.

Real time PCR machine

The hardware

The instrument consists of a thermal cycler, light source, interchangeable light filters and a sensitive light recording device. The light source is either halogen lamp or LED and the recording devices may be CCD camera, photodiode, or photomultiplier tube. The machines vary in the capability of detecting the number of colours. A basic machine can detect two colours while more advanced machines can detect up to five or more colours. The choice of more colour detection is useful for simultaneous detection of greater number of probes with different fluorochromes in a single reaction tube for example multiplex PCR.

The software

The instrument software provides an interface for entering thermal cycling parameters and choice of dye (fluorochrome) and quencher selection. As the run starts the real time graphic record of amplification after each cycle is displayed on the screen. At the end of the run the amplification record can be seen on linear or logarithmic scale.

Threshold cycle (Ct)

The first significant increase in the amount of PCR product as seen in the real time graphic recording of a PCR is called "Cycle-threshold" or "Ct". It correlates indirectly with the starting amount of template DNA. Higher the amount of template DNA smaller is the Ct (Fig. 6.3). The real time PCR software provides options for calculating Ct automatically or manually.

Rn

Rn is an indicator of the amount of fluorescence generated in a PCR. The amount of fluorescence (Rn) of a test reaction and a non target control (NTC) are called Rn+ and Rn- respectively. A difference between the Rn+ and R- is called ΔRn. The Ct values are derived from plotting the cycle numbers against the ΔRn (Fig. 6.3).

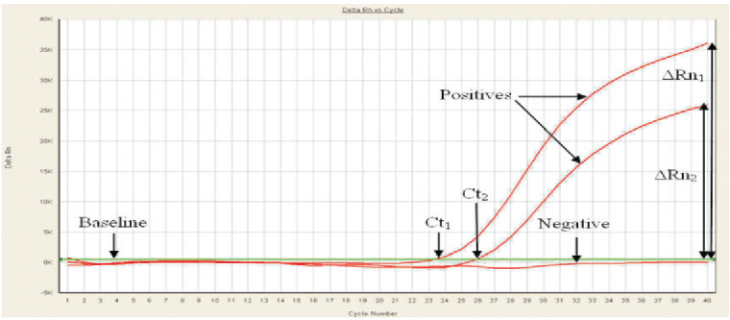


Fig. 6.3. Real time PCR plot on a linear scale. The cycle numbers are shown on X-axis while the amount of fluorescence is shown on Y-axis. The Ct1 (23.6) and Ct2 (26.0) indicate that there is greater quantity of target DNA in the sample 1 than in the sample 2. This is also reflected in the ΔRn values of the two samples.

Linear versus logarithmic recording

In PCR one molecule of DNA is doubled with every cycle (Table 6.2). In the cycle numbers 15-20, although the DNA molecules are doubling in number, their total number is far less as compared to the number of DNA molecules produced in the later cycles. In a real time PCR our main interest is in knowing the first detectable change in the amount of fluorescence (increase in DNA). If the result of real time PCR is plotted on a linear scale it becomes difficult to appreciate the smaller change in fluorescent signal in the earlier cycles. However, if the results are plotted on a log scale the change in the earlier cycles becomes more pronounced and easy to recognize. The software of the real time PCR machines provides option for displaying the results in linear as well as log scale.

Table 6.2. The number of DNA molecules produced with each cycle of PCR from one molecule of DNA. The increase in the number of DNA molecules in the earlier cycles (10-19) is much less than that observed in the cycles later than 20.

Cycle No.	DNA Molecules	Cycle No.	DNA Molecules	Cycle No.	DNA Molecules
1	2	11	2048	21	2,097,152
2	4	12	4096	22	4,194,304
3	8	13	8192	23	8,388,608
4	16	14	16,384	24	16,777,216
5	32	15	32,768	25	33,554,432
6	64	16	65,532	26	67,108,864
7	128	17	131,072	27	134,217,728
8	256	18	262,144	28	268,435,456
9	512	19	524,288	29	536,870,912
10	1024	20	1,048,576	30	1,073,741,824

Real time PCR applications

The best use of real time PCR is in quantification of DNA and RNA. The quantification may be absolute, relative or comparative. In absolute quantification accurately quantified standards are used to make a standard curve. In relative quantification serial dilutions of a calibrator are used to make a standard curve. In comparative quantification the calibrator is used only once and the calculations are made mathematically.

Another common use of real time PCR is in detection of pathogens like viruses or bacteria. Here the major advantage of real time PCR is in elimination of end point processing which is a major source of false positives due to carry over of amplified product from one sample to another. An additional advantage is in the reduction of steps in analysis and hence less chances of errors.

Real time PCR may also be used in multiplex allelic discrimination and single nucleotide polymorphism (SNP) assays. Multiple TaqMan® probes with different fluorescent dyes can be used for different alleles. In SYBR green based methods melting curve analysis can be used to detect multiple amplified fragments.

PCR efficiency

Efficiency of a PCR means the rate at which the amplified product is generated. Real time PCR may be used to calculate the efficiency of amplification. The efficiency is usually gauged by the number of cycles in which a given amount of DNA is increased ten folds (times). In a 100% efficient PCR 10 fold increase in the target DNA is achieved in 3.32 cycles. A value higher than this means less efficient PCR and vice versa. The efficiency of PCR is measured from the slope of the serially diluted standard curve. A steep slope indicates low efficiency whereas a gradual slope indicates a more efficient PCR. The PCR efficiency is dependent on a large number of variables including the PCR machine and the reaction conditions. Out of the later the composition of the PCR mix is a major determinant of the PCR efficiency. Several additives that enhance PCR (Chapter 3) can improve PCR efficiency. PCR efficiency between 90-110% is considered acceptable. A higher efficiency of PCR is more likely to detect smaller amounts of the target.

R^2

In a real time PCR there is a positive correlation between Ct value and the concentration of DNA. Under perfect conditions the Ct values of ten fold dilution of DNA standard when plotted against the concentration should show a straight line (100% correlation or R^2=1.0). However, inconsistencies in carrying out PCR, notably the pipetting errors, may cause varying degrees of deviation of the individual Ct values from the straight line. This may result in R^2 values below 1.0.

Setting up a real time PCR

A real time PCR for Mycobacterium tuberculosis is described as an example. The PCR primers and the TaqMan® probe (Fig. 6.4) for the insertion sequence element IS986/IS6110 of Mycobacterium tuberculosis CDC1551 (GenBank accession: AE000516.2) were designed by the Primer Express® software (Applied Biosystems).

Fig. 6.4. The primers and the TaqMan® probe for Mycobacterium tuberculosis insertion sequence (IS6110): (GenBank: AE000516.2).

CGGAGCTGCGCGATGGCGAACTCAAGGAGCACATCAGCCGCGTCCACGCCGCCAAC
TACGGTGTTTACGGTGCCCGCAAAGTGTGGCTAACCCT GAACCGTGAGGCATCGAG
GTGGCCAGATGCACCGTCGAACGGCTGATGACCAAACTCGGCCTGTCCGGGACCAC
CCGCGGCAAAGCCCGCAGGACCACGATCGCTGATCCGGCCACAGCCCGTCCCGCCG

Forward primer: 5'-GAACCGTGAGGGCATCGA
Reverse primer: 5'-ACAGGCCGAGTTTGGTCATC
TaqMan® Probe: 6 FAM 5'-CCAGATGCACCGTCGAACGGC-BHQ1
Amplified product: 64bp

In this example Positive Standard DNA of Mycobacterium tuberculosis (AcuGen, USA) was used that contains 500,000 copies of DNA per micro liter.

1. Make primer/probe mix by diluting the two primers and the probe at 5pmol/µl each (Chapter 3) in a single tube.
2. Label 4 x 0.2 ml PCR reaction tubes in a rack.

3. Take two 0.2 ml tubes and label them as 1/10 and 1/100 and add 45µl of DEPC treated water in each tube.
4. Add 5µl of the standard DNA to the tube labeled 1/10 and mix. Transfer 50µl from this tube to the tube labeled 1/100 and mix. This will give 50,000/µl and 5000/µl standard DNA respectively in each dilution.
5. Make 80µl real time PCR master mix for 4 reactions in 0.2 ml tube by mixing the following:

    - PCR buffer (Chapter 3): 80µl
    - Primer/probe mix: 4µl
    - Taq polymerase (5 units/µl) 0.4µl (0.1µl/reaction)

6. Transfer 21µl master mix to each of the four labeled tubes.
7. Add 5ul of the undiluted, 1/10 and 1/100 diluted DNA standard and DEPC treated water (no target control) to each of the four labeled tubes.
8. Firmly close the lids of the tubes and put these in the real time thermal cycler.
9. Set the following thermal cycling parameters:

    a. Initial denaturation 95°C for 5 minutes.
    b. Forty cycles each of denaturation at 95°C for 15 seconds, and annealing/extension at 60°C for 1 minute.

10. Set instrument software parameters as given in the instruction manual. Choose FAM as reporter.

Result:

The result of the real time PCR for the ten fold dilutions of Mycobacterium tuberculosis control DNA run on Rotor Gene 6000 is shown in Fig 6.5. The Ct values of each dilution are shown in Table. 6.3.

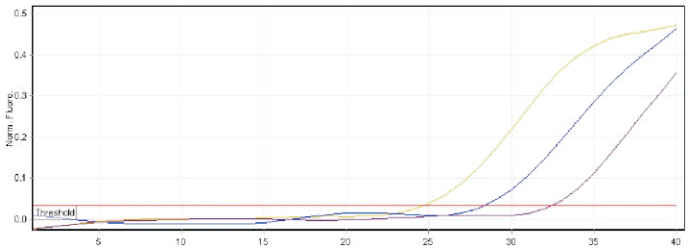


Fig. 6.5. Real time PCR results of 10 fold diluted positive control DNA of Mycobacterium tuberculosis. The Ct of the first dilution is the lowest and it increases with each dilution.

Table. 6.3. Ct values of the ten fold diluted control DNA of Mycobacterium tuberculosis.

Dilution	Copy number/µl	Ct
Undiluted Standard DNA	500,000	25.12
1/10 diluted Standard DNA	50,000	28.49
1/100 diluted Standard DNA	5,000	32.53
No target control	0	39.36

Absolute quantification of DNA by real time PCR

The above example of ten fold dilutions of the standard DNA can be used to determine the DNA copy number in one or more unknown samples. The samples with unknown concentration of DNA are also run in a similar way as the ten fold diluted DNA was run. The absolute quantification is done as follows:

a. Use the software of the real time PCR machine to plot the standard curve of the copy numbers and the respective Ct values (Fig. 6.6).
b. The software calculates the concentration (copy number) in the unknown sample(s).

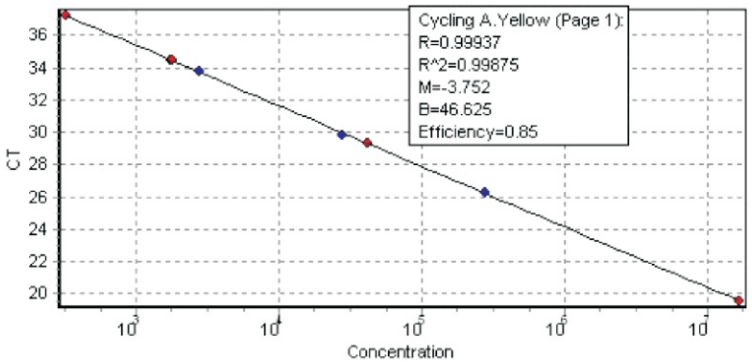


Fig. 6.6. Standard curve plot between the DNA concentration (x-axis) and the corresponding Ct value (y-axis) of the 10 fold DNA dilutions (blue dots) and unknown DNA samples (red dots) run on the Rotor Gene 6000 machine. The inside window also displays the related PCR statistics.

Relative Quantification of DNA by real time PCR

The amount of target DNA in a sample can also be calculated from the difference in the Ct values of a target DNA and an internal reference (internal control DNA) or a DNA standard with known concentration. The relative quantification is subject to the quality of the RNA or DNA in the target and the reference material. Ideally the quality of RNA or DNA of the two should be similar. The target and the reference RNA or DNA can be run in the same or two different reaction tubes. The difference in the Ct value between the unknown (target) and the standard (reference) is called ΔCt.

ΔCt = Ct target – Ct reference (standard)
The amount of target is given by the mathematical formula $2^{-\Delta Ct}$
For example at ΔCt of +2.0 the amount of target would be:
$2^{2.0} = 4.0$
Or the target sample has four times (fold) DNA or RNA as compared to that present in the reference (standard).
At ΔCt of -2.0 the amount of target would be:
$2^{-2.0} = 0.25$

Or the target sample has 0.25 (1/4) fold DNA or RNA as compared to that present in the reference (standard).
A scientific calculator or Excel spreadsheet can be used to calculate complex values of ΔCt like 2.7 or -3.4 etc.

DNA quantification by real time PCR versus gel electrophoresis

Fig. 6.7 shows the real time PCR plot of a serially diluted DNA sample. All dilutions have clearly distinct Ct values that correlate with the amount of DNA in the sample. It is also seen that after the 30th cycles the reaction enters a platue phase. If gel electrophoresis is done on the same samples at the end of 35 cycles they would show bands of only slightly differing intensities. This is why real time PCR is considered superior to gel electrophoresis for DNA quantification by PCR.

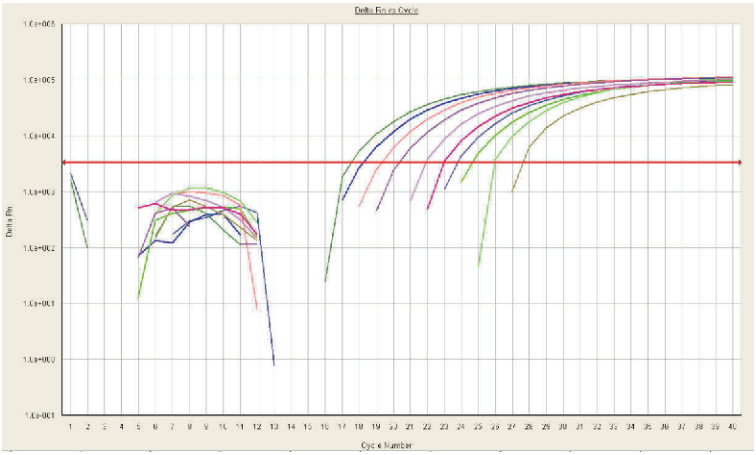


Fig. 6.7. Real time plot of a serially diluted sample of DNA. All dilutions have clearly distinct Ct values that correlate with the amount of DNA in the sample. After the 30th cycles the reaction ha entered a platue phase. If gel electrophoresis is done on the same samples at the end of 35 cycles they would show bands of only slightly differing intensities.

Bibliography

1. Dorak MT (ed) (2006) Real Time PCR. Taylor & Francis Group, New York.
2. Ponchel F, Toomes C, Bransfield K, Leong FT, Douglas SH, Field SL, Bell
3. SM, Combaret V, Puisieux A, Mighell AJ, Robinson PA, Inglehearn CF, Isaacs JD, Markham AF (2003) Real-time PCR based on SYBR-Green I fluorescence: An alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions. BMC Biotechnology 3: 18-31.

7. Amplification Refractory Mutation System (ARMS)

The Amplification Refractory Mutation System (ARMS) is an application of PCR in which DNA is amplified by allele specific primers. In PCR mismatch at the 3' end of the primer can dramatically reduce the annealing and hence the amplification. This is due to the absence of 3' to 5' exonuclease proofreading activity of Taq polymerase. High fidelity DNA polymerases, that have this activity, cannot be used in ARMS. It is an extremely useful method for identification of point mutations or polymorphisms.

Since the ARMS PCR is mostly done to identify a mutation or a polymorphism it is also important that it should be able to identify whether the change in DNA is heterozygous or homozygous. A heterozygote or homozygote is differentiated by using ARMS primers for the mutant/polymorphic and the normal (wild type) alleles. The reactions for the mutant and the normal alleles are usually carried out in separate tubes. But these may be done in the same tube after labeling the two primers with different fluorescent dyes.

ARMS primer design

General principles of designing a PCR primer as discussed in chapter 3 also apply to the ARMS primers. The ARMS PCR requires a pair of primers including a common and an ARMS primer. The common primer is like any other PCR primer. But the ARMS primer has the following special features:

1. The primer is usually 30 bases in length.
2. The nucleotide at the 3' end of the primer should be complementary to the target nucleotide i.e. G for C or C for G and T for A or A for T. Mismatch at this position can dramatically reduce the amplification. A:G, G:A, and C:C mismatches have the worst effect whereas the other mismatches have varying degrees of effect. For example in a mutation with A-T substitution the ARMS primer for the mutant allele should have the last nucleotide complementary to the nucleotide T i.e. it should have A. The primer for the normal allele at the same position should be complementary to the nucleotide A i.e. it should have T (Fig 7.1).
3. An additional mismatch at one of the last five nucleotides of the ARMS primer further increases its specificity.
4. It is customary to include an internal PCR control in ARMS reactions. A pair of primers is designed to amplify a region of the gene of interest that usually is free of mutations. An amplification of the internal control region and no amplification by the ARMS primer indicate a true negative. In a false negative result neither the ARMS primer nor the internal control shows any amplification. There could be several reasons for the false negative result e.g. too little or too much DNA, poor quality of DNA template, failure to add primer, Taq, or other reagents and presence of PCR inhibitors.

5. The sensitivity and specificity of an ARMS reaction can be controlled by stringent reaction conditions. Good primer design, higher annealing temperature and limited number of cycles are important in avoiding false results. The number of cycles should be just enough to give a clear positive result. Increasing the number of cycles un-necessarily can cause false positives. The usual length of ARMS primer is 30 bases. Primers of this length have a high Tm and annealing temperature and are therefore more specific.

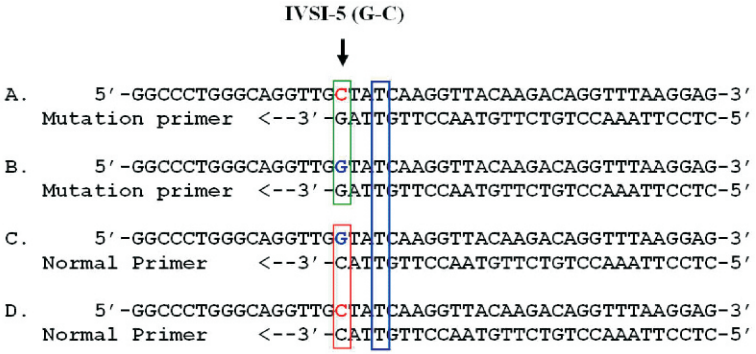


Fig. 7.1 ARMS primers for the normal and the mutant allele of a point mutation, IVSI-5 (G-C), in the β-globin gene. Segment A shows the matched ARMS primer for the mutation (C), segment B shows the mismatch between the mutant ARMS primer and the normal sequence (G), segment C shows the matched ARMS primer for the normal sequence (G), segment D shows the mismatch between the normal ARMS primer and the mutation (C). A deliberate mismatch (T:T) is also added at position enclosed by blue box.

PCR conditions for ARMS

Locus:	β-globin gene
GenBank accession	NG_000007.3
Allele:	β-thalassaemia mutation IVSI-5 (G-C)

- Forward primer: 5'-ACCTCACCCTGTGGAGCCAC
- Reverse primer (ARMS)
  5'-CTCCTTAAACCTGTCTTGTAACCTTGTTAG
- Amplified product 285 bp
- Control primer (forward):
  5'-CAATGTATCATGCCTCTTTGCACC
- Control primer (reverse):
  5'-GAGTCAAGGCTGAGAGATGCAGGA
- Amplified product: 861 bp

•

- Reaction volume: 25µl
- PCR mix: 22µl
- Primer concentration: 1µl (5 pmol each/µl) (Chapter 3)
- Taq polymerase: 0.5 units (0.1µl)
- Template DNA: 2µl (~300ng)

Thermal cycling:

- Initial denaturation: 1 minute at 94°C
- No. of cycles: 25
- Denaturation: 1 minute at 94°C
- Annealing: 1 minute at 65°C
- Extension: 1 minute 30 seconds at 72°C
- Final extension: 3 minute at 72°C

Electrophoresis: 10 X 10 cm 6% polyacrylamide, 40 minutes at 150 volts.
Staining: 0.1% silver nitrate
Result: Fig. 7.2

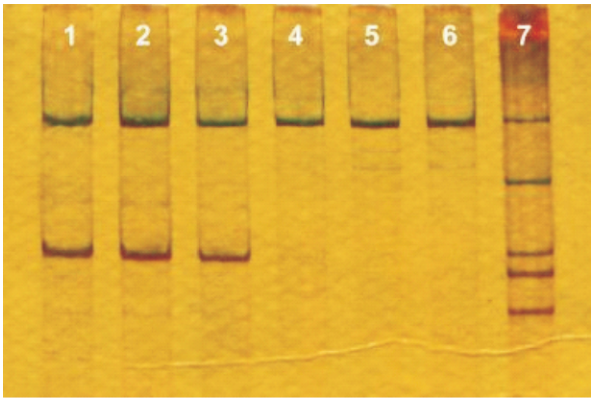


Fig. 7.2. Silver stained polyacrylamide gel electrophoresis after ARMS PCR. All lanes show 861bp internal control fragment. Lanes 1-3 shows 285bp fragment of IVS1-5 mutation. Lanes 4-5 are negative for the same mutations. Lane 7 shows allelic ladder for various thalassaemia mutations.

Bibliography

1. Kwok S, Kellogg DE, McKinney N, Spasic D, Goda L, Levenson C, Sininsky JJ (1989) Effect of primer template mismatches on the polymerase chain reaction: Human immunedeficiency virus type 1 model studies. Nucleic Acids Res 18: 999-1005.
2. Newton CP, Graham A, Heptinstall LE, Powell SJ, Summers C, Kalsheker N, Smith JC, Markham AF (1989) Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). Nucleic Acids Res 17: 2503-2516.
3. Old JM, Varawalla NY, Weatherall DJ (1990) Rapid detection and prenatal diagnosis of -thalassaemia: studies in Indian and Cypriot populations in the UK. Lancet 336: 834-837.

8. Restriction Fragment Length Polymorphism (RFLP)

Restriction endonucleases are enzymes of bacterial origin that can cut DNA at specific sequences called “restriction sites”. As a result of digestion by a restriction enzyme the DNA sequence is cut in to fragments of varying lengths. This forms the basis of restriction fragment length polymorphism (RFLP). It is a very useful technique in identification of single nucleotide polymorphisms (SNP) or point mutations causing disease.

The RFLP can detect SNP or point mutations when the latter create or abolish a restriction site. The presence of a restriction site is designated as “+” and its absence as “-“. The sequence containing the SNP or the mutation is amplified by a pair of primers and the amplified DNA is incubated with the restriction enzyme. If the SNP is present the amplified DNA is cut in to two pieces at a point where the SNP is present (+). If the SNP is not present the DNA fragment remains as single piece (-). The case would be reversed if the SNP abolishes the restriction site. A heterozygote of SNP would have the cut and the uncut fragments (-/+). The homozygote would have only the cut (+/+) or the uncut (-/-) fragments.

Example 1 (SNP)

DNA sequence in the 5' un-translated region of the β-globin gene contains C-T polymorphism (GenBank NG_000007.3). The two alleles of the SNP are highlighted in Fig. 8.1. The polymorphism can be identified by digestion with restriction enzyme Rsa-I that can cut the GTAC sequence (Fig 8.2).

Allele 1 (GTAC)

```
5′-ATATATGTATATGTATGTGTGTACATATACACACATATATATATATATT-3′
3′-TATATACATATACATACACACATGTATATGTGTGTATATATATATATAA-5′
```

Allele 2 (GTAT)

```
5′-ATATATGTATATGTATGTGTGTATATATACACACATATATATATATATT-3′
3′-TATATACATATACATACACACATATATATGTGTGTATATATATATATAA-5′
```

Fig. 8.1. A heterozygous C-T polymorphism is shown in the red box.

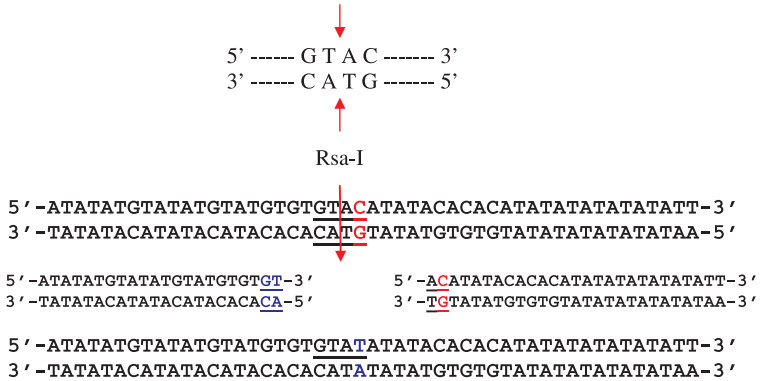


Fig. 8.2. DNA containing the GTAC sequence is cut by restriction enzyme Rsa-I into two pieces. The GTAT allele would not be cut by Rsa-I and therefore would remain as a single piece.

PCR protocol for RFLP

A 213bp fragment of DNA containing the C-T SNP is amplified by a pair of primers (Fig. 8.3). The amplified DNA is incubated overnight at 37°C with the restriction enzyme Rsa-I. The C-T polymorphism when present results in cutting of the 213bp fragment in to two pieces of 66bp and 147bp each.

TTTAGACATAATTTATTA**GCATGCATGAGCAAATTAAGA**AAAACAACAACAAATGAA
TGCATATATATGTATATGTATGTGTGTA**[C/T]**ATATACACACATATATATATATAT
TTTTTCTTTTCTTACCAGAAGGTTTTAATCCAAATAAGGAGAAGATATGCTTAGAAC
CGAGGTA**GAGTTTTCATCCATTCTGTCCTG**TAAGTATTTTGCATATTCTGGAGACGC

Fig. 8.3. PCR primers for the C-T polymorphism in the 5' untranslated region of β-globin gene.

- Locus: β-globin gene
- GenBank accession: NG_000007.3
- Forward primer: 5'-GCATGCATGAGCAAATTAAGA
- Reverse primer: 5'-TCTTCCTGCGTCTCCAGAAT
- Amplified product: 213 bp
- Restriction enzyme: Rsa-I
- Restriction fragments: 66bp and 147bp
- Reaction volume: 25µl
- Primer concentration: 1µl (5 pmol each/µl)
- Taq polymerase: 0.5 units (0.1µl)
- Template DNA: 1µl (~300ng)
- Thermal cycling:
  - Initial denaturation: 1 minute at 94°C
  - No. of cycles: 30
  - Denaturation: 1 minute at 94°C
    - Annealing: 1 minute at 57°C
    - Extension: 1 minute 30 seconds at 72°C

- Final extension: 3 minute at 72°C
- Restriction enzyme (Rsa-I): 1µl
- Amplified product: 15 µl
- Incubation: Overnight at 37°C
- Electrophoresis: 10 X 10 cm 6% polyacrylamide
  Load 3µl digested amplified product in 3µl loading dye
  Run for 20 minutes at 150 volts
- Staining: 0.1% silver nitrate

Example 2

Sickle haemoglobin is caused by A-T substitution in Codon 6 of β-globin gene. The mutation can be identified by RFLP using the restriction enzyme Dde-I. The enzyme recognizes the sequence CTNAG where N stands for any nucleotide (Fig. 8.4).

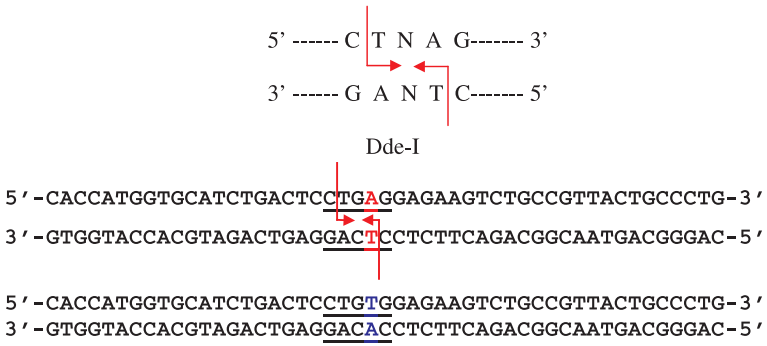


Fig. 8.4 The recognition site of restriction enzyme Dde-I can cut the DNA containing A-T substitution causing Hb-S mutation in the β-globin gene.

TAGACCTCACCCTGTGGAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGG
CAGGAGCCAGGGCTGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACATTTGCTTCTGA
CACAACTGTGTTCACTAGCAACCTCAAACAGACACCATGGTGCATCTGACTCCTG[A/T]GGA
GAAGTCTGCCGTTACTGCCCTGTGGGGCAAGGTGAACGTGGATGAAGTTGGTGGTGAGGCCCT
GGGCAGGTTGGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCATGTGG
AGACAGAGAAGACTCTTGGGTTTCTGATAGGCACTGACTCTCTCTGCCTATTGGTCTATTTTC
CCACCCTTAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAGTCCTTTGGGGATC
TGTCCACTCCTGATGCTGTTATGGGCAACCCTAAGGTGAAGGCTCATGGCAAGAAAGTGCTCG
GTGCCTTTAGTGATGGCCTGGCTCACCTGGACAACCTCAAGGGCACCTTTGCCACACTGAGTG
AGCTGCACTGTGACAAGCTGCACGTGGATCCTGAGAACTTCAGGGTGAGTCTATGGGACGCTT
GATGTTTTCTTTCCCCTTCTTTTCTATGGTTAAGTTCATGTCATAGGAAGGGGAT

Fig. 8.5. The primer pair for amplification of A-T substitution in codon 6 of β-globin gene. The 597bp amplified product contains three restriction sites recognized by Dde-I at positions +145, +523 and +560 from the forward primer (red). The first of the three restriction sites is abolished by the presence of A-T substitution (Hb-S).

The 597bp fragment of β-globin gene containing the A-T substitution (Fig. 8.5) is amplified by the following pair of primers:

- Forward Primer: 5'- GGCCAATCTACTCCCAGGAG
- Reverse Primer: 5'-ACATCAAGCGTCCCATAGAC

The amplified fragment contains three restriction sites identified by the enzyme Dde-I. The first site, present at 145bp downstream from the forward primer, is altered when Hb-S mutation is present. The other two sites serve as internal controls for the restriction enzyme Dde-I. The net result of Dde-I digestion of the 597bp fragment is generation of different patterns of restriction fragments when the Hb-S mutation is present or absent.

PCR Protocol:

- Reaction volume: 25µl
- Primer concentration: 1µl (5 pmol each/µl)
- Taq polymerase: 0.5 units (0.1µl)
- Template DNA: 1µl (~300ng)
- Thermal cycling:
  - Initial denaturation: 1 minute at 94°C
  - No. of cycles: 30
    - Denaturation: 1 minute at 94°C
    - Annealing: 1 minute at 57°C
    - Extension: 1 minute 30 seconds at 72°C
  - Final extension: 3 minute at 72°C
- Restriction enzyme (Dde-I): 1µl
- Amplified product: 15 µl
- Incubation: Overnight at 37°C
- Electrophoresis: 10 X 10 cm 6% polyacrylamide
  Load 3µl digested amplified product in 3µl loading dye
  Run for 20 minutes at 150 volts
- Staining: 0.1% silver nitrate

Examples 3

HFE gene mutation (H63D)

H63D mutation in the HFE gene on chromosome 6 can be detected by RFLP. The restriction enzyme Mbo-1 can cut the DNA sequence NGATCN:

```
5'...G A T C...3'
3'...C T A G...5'
```

A 208bp fragment of the DNA sequence containing the H63D mutation (GenBank NT_007592.15) is amplified by the following primers:

H63D-F	5'-ACATGGTTAAGGCCTGTTGC
H63D-R	5'-GCCACATCTGGCTTGAAATT

The normal (wild type) DNA is cut by the restriction enzyme (Mbo-1) resulting in two fragments of 138 and 70bp. The presence of H63D mutation abolishes the restriction site. In a heterozygote (-/+) the uncut 208bp fragment and the cut 138bp and 70bp fragments are present. Homozygotes (-/-) for the mutation show only the 208bp fragment where as the homozygotes for the normal allele (+/+) show two fragments of 138bp and 70bp (Fig. 8.6).

PCR Protocol:

- Reaction volume: 25µl
- Primer concentration: 1µl (5 pmol each/µl)
- Taq polymerase: 0.5 units (0.1µl)
- Template DNA: 1µl (~300ng)
- Thermal cycling:
  - Initial denaturation: 1 minute at 94°C
  - No. of cycles: 30
    - Denaturation: 1 minute at 94°C
    - Annealing: 1 minute at 57°C
    - Extension: 1 minute 30 seconds at 72°C
  - Final extension: 3 minute at 72°C
- Restriction enzyme (Dde-I): 1µl
- Amplified product: 15 µl
- Incubation: Overnight at 37°C
- Electrophoresis: 10 X 10 cm 6% polyacrylamide
  Load 3µl digested amplified product in 3µl loading dye
  Run for 20 minutes at 150 volts
- Staining: 0.1% silver nitrate

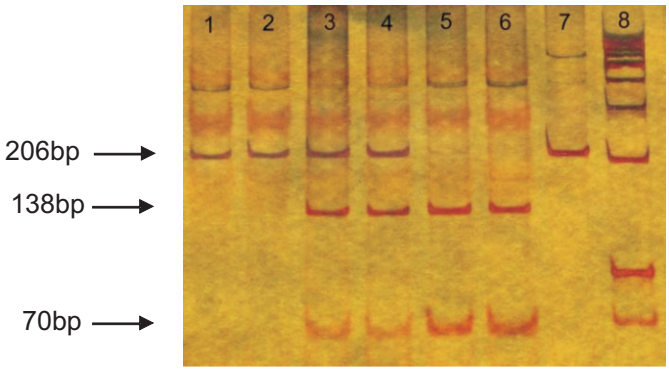


Fig. 8.6. Silver stained polyacrylamide gel electrophoresis of the Mbo-I restriction enzyme digested fragments of the HFE gene identifying the H63D mutation. The -/- (lanes 1 & 2) indicates homozygous and -/+ (lanes 3 & 4) indicates heterozygous for the mutation. The +/+ (lane 5 & 6) indicates normal genotype. Lane 7 shows 206bp uncut fragment.

Bibliography

1. Geever RF, Wilson LB, Nallaseth FS, Milner PF, Bittner M, Wilson JT (1981) Direct identification of sickle cell anemia by blot hybridization. Proc Natl Acad Sci USA 8: 5081-5085.
2. Old JM, Heath C, Fitches A, Thein SL, Jeffreys AJ, Petrou M, Modell B, Weatherall DJ (1986) Meiotic recombination between two polymorphic restriction sites within the globin gene cluster. J Med Genet 23: 14-18.
3. Varawalla NY, Fitchies AC and Old JM (1992) Analysis of -globin haplotypes in Asian Indians: origin and spread of -thalassaemia on the Indian subcontinent. Hum Genet 90: 443-449.
4. Tsui WMS, Lam PWY, Lee KC, Ma KF, Chan YK, Wong MWY, Yip SP, Wong CSC, Chow ASF, Lo STH (2000) The C282Y mutation of the HFE gene is not found in Chinese haemochromatotic patients: multicentre retrospective study. HKMJ 6: 153-8.

9. Short Tandem Repeat (STR) Analysis

Short tandem repeats (STR) are randomly distributed DNA sequences in which 2-6bp are tandemly repeated. These are scattered on all chromosomes including the autosomes as well as the sex chromosomes. An example of four base pair (tetranucleotide) repeat STR on chromosome 5 is shown in Fig. 9.1 The sequence "GATA" is repeated several times. The numbers of STR repeats vary between individuals and the alleles are named according to the number of repeat units these may contain. Every individual inherits an allele each from its parents. The two alleles may be identical (homozygote) or different (heterozygote or more precisely compound heterozygote).

ATATTTTAATAGCAAGTATGTGACAAGGGTGATTTTCCTCTTTGGTATCCTTATGTA
ATATTTTGAAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGGT
AGATAGAGGTATAAATAAGGATACAGATATAGNTACAAATGTTGTAAACTGTGGCTA
TGATTGGAATCACTTGGCTAAAAAGCGCTNAAGCNTTCCTCTGNGAGAGGCAATTAC

Fig. 9.1 GATA repeats at the D5S818 locus (GenBank: G08446.1) The sequence GATA is repeated 11 times.

PCR amplification of STRs

STRs can be amplified by a pair of primers flanking the tandem repeats. The resulting amplified products vary in length depending on the number of repeat units present in the allele (Fig. 9.2A). In a compound heterozygote with different number of repeats on the two chromosomes, two different sized amplified products are formed. In a homozygote with the same number of repeats on the two chromosomes two amplified fragments of the same size are produced. The amplified fragments vary in length depending on the number of repeats in the allele. If at a tetranucleotide repeat locus one allele has 5 repeats of 4bp each and the other allele has 8 repeats the two amplified fragments would differ by 12bp (4 + 4 + 4).

Polyacrylamide gel electrophoresis

The amplified fragments of STRs at any given locus are only a few base pairs different. Such small differences are best resolved by Polyacrylamide gel electrophoresis. After PCR amplification the STRs are analysed by relatively long polyacrylamide gels. The gels are usually run for several hours. Various alleles are recognized by simultaneously running accurate DNA size standards or allelic ladders (Fig. 9.2B).

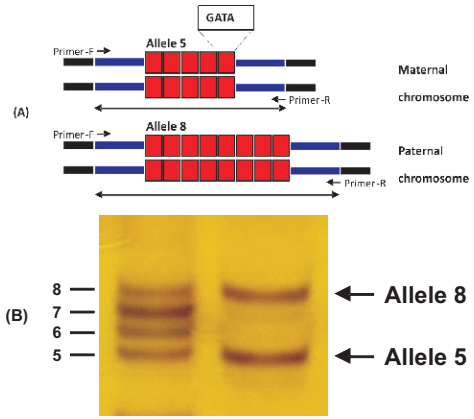


Fig. 9.2. (A) PCR amplification of STR locus D5S818 having "GATA" repeats. The tandem repeats on the maternal and the paternal chromosomes are amplified by a pair of primers (F & R) flanking the repeat sequences.

(B) Polyacrylamide gel electrophoresis of the amplified product and allelic ladder shows the presence of allele 5 and 8.

Genetic analyzer

The manual sizing of STR alleles is laborious and time consuming. Automated genetic analyzer provides an extremely efficient but expensive alternate to the manual method. It uses a long thin bored capillary for electrophoresis of amplified DNA (Chapter 5). The STRs are amplified by fluorescent labeled primers and are run with a fluorescent labeled DNA size standard in the same capillary. The genetic analyzer detects the amplified products by laser excitation of the fluorescent dyes. It is an extremely sensitive and accurate method of STR allele sizing. Very small amounts of target DNA (<0.5ng) can be detected. Its other advantages include high speed and analysis of multiple amplified products labeled with different fluorescent dyes in the same tube. Commercial kits with amplification of up to fifteen different STR loci in the same tube are also available.
STR analysis by manual method

Locus:	D5S818 (Fig 9.3)
GenBank accession:	G08446.1
Forward primer:	5'-AAGGGTGATTTTCCTCTTTGG
Reverse primer:	5'- AGCGCTTTTTAGCCAAGTGA
Amplified product:	136-176 bp (1-11 repeats)

TCTAATTAAAGTGGTGTCCCAGATAATCTGTACTAATAAAAGTATATTTTAATAG
CAAGTATGTGAC**AAGGGTGATTTTCCTCTTTGG**TATCCTTATGTAATATTTTGAA
GATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGGTAGATAGAG
GTATAAATAAGGATACAGATATAGNTACAAATGTTGTAAACTGTGGCTATGATTG
GAA**TCACTTGGCTAAAAAGCGCT**NAAGCNTTCCTCTGNGAGAGGCAATTACTTTT
TTNCTTAGGNACTNCCTCANCAGTCTNTTNGC

Fig. 9.3. PCR primers for amplification of the GATA repeat at D5S818 locus (GenBank accession: G08446.1).

Reaction conditions

- Reaction volume: 25 µl
- Primer concentration: 1 µl (5 pmol each/µl)
- Taq polymerase: 0.5 units (0.1µl)
- Template DNA: 1µl (200ng)
- Thermal cycling:
  - Initial denaturation: 1 minute at 94°C
  - No. of cycles: 30
- Denaturation: 48 seconds at 93°C
- Annealing: 48 seconds at 60°C
- Extension: 1 minute at 72°C
- Final extension: 3 minute at 72°C
- Electrophoresis: 16 X 20 X 0.1cm 6% polyacrylamide gel. Load 3µl amplified product in 3µl loading dye Include 3µl allelic ladder in 3µl loading dye Run for 5 hours at 200 volts.
- Staining: 0.1% Silver nitrate.
- Result: Fig 9.4

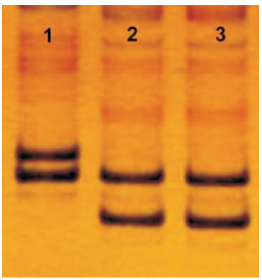


Fig. 9.4. Polyacrylamide gel electrophoresis of the PCR products of amplification at D5S818 locus.

Preparation of allelic ladder for STR analysis by gel electrophoresis

1. Select DNA samples of known genotypes at different loci previously analyzed on genetic analyzer. The selected samples should represent the alleles usually found in the target population.
2. Amplify each DNA sample in multiples of five at the locus for which the ladder is required. Use the amplification protocol described above.
3. Pool the multiply amplified products of all the samples.
4. The amplified DNA is a potent source of cross contamination therefore only use pipettes marked as “amplified DNA only”.
5. Mix the pooled amplified products and make aliquots of 100µl.
6. Store at -20°C.
7. Use 3µl of the pooled amplified products (allelic ladder) for each electrophoresis run (Fig. 9.4).

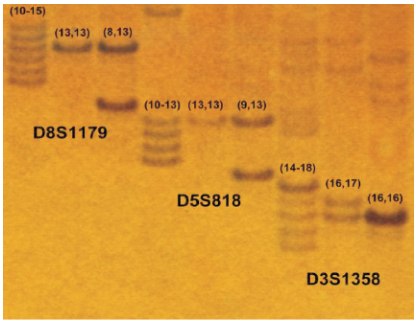


Fig. 9.5 Silver stained polyacrylamide gel electrophoresis of PCR amplified products at D8S1179, D5S818 and D3S1358 loci and the allelic ladders. The allele sizes in each allelic ladder and two unknown samples are shown in brackets.

PCR primers for amplification of STR loci

The lists of primers for amplification of autosomal and Y-chromosomal STRs are given in Tables 9.1 & 9.2.

Table 9.1. PCR primers for CODIS core STR Loci.

Locus	Primer sequence
D3S1358-F	5'-ACTGCAGTCCAATCTGGGT
D3S1358-R	5'-ATGAAATCAACAGAGGCTTG
D5S818-F	5'-GGGTGATTTTCCTCTTTGGT
D5S818-R	5'-TGATTCCAATCATAGCCACA
D7S820-F	5'-TGTCATAGTTTAGAACGAACTAACG
D7S820-R	5'-CTGAGGTATCAAAAACTCAGAGG
D8S1179-F	5'-TTTTTGTATTTCATGTGTACATTCG
D8S1179-R	5'-CGTAGCTATAATTAGTTCATTTTCA
D13S317-F	5'-ACAGAAGTCTGGGATGTGGA
D13S317-R	5'-GCCCAAAAAGACAGACAGAA
D16S539-F	5'-GATCCCAAGCTCTTCCTCTT
D16S539-R	5'-ACGTTTGTGTGTGCATCTGT
D18S51-F	5'-CAAACCCGACTACCAGCAAC
D18S51-R	5'-GAGCCATGTTCATGCCACTG
FGA-F	5'-GCCCCATAGGTTTTGAACTCA
FGA-R	5'-TGATTTGTCTGTAATTGCCAGC
TPOX-F	5'-CACTAGCACCCAGAACCGTC
TPOX-R	5'-CCTTGTCAGCGTTTATTTGCC
CSF1PO-F	5'-AACCTGAGTCTGCCAAGGACTAGC
CSF1PO-R	5'-TTCCACACACCACTGGCCATCTTC
TH01-F	5'-GTGGGCTGAAAAGCTCCCGATTAT
TH01-R	5'-ATTCAAAGGGTATCTGGGCTCTGG
vWA-F	5'-CCCTAGTGGATAAGAATAATC
vWA-R	5'-GGACAGATGATAAATACATAGGATGGATGG
AmgI-F	5'-ACCTCATCCTGGGCACCCTGG
AmgI-R	5'-AGGCTTGAGGCCAACCATCAG

Table: 9.2. PCR primers for Y-STR loci.

Locus	Primer sequence
DYS 19-F	5'-CTACTGAGTTTCTGTTATAGT
DYS 19-R	5'-ATGGCCATGTAGTGAGGACA
DYS 385a/b-F	5'-AGCATGGGTGACAGAGCTA
DYS 385a/b-R	5'-TGGGATGCTAGGTAAAGCTG
DYS 389-I/II -F	5'-CCAACTCTCATCTGTATTATCTAT
DYS 389-I/II -R	5'-TCTTATCTCCACCCAGA
DYS 390-F	5'-TATATTTTACACATTTTTGGGCC
DYS 390-R	5'-TGACAGTAAAATGAACACATTGC
DYS 391-F	5'-CTATTCATTCAATCATACACCCA
DYS 391-R	5'-GATTCTTTGTGGTGGGTCTG
DYS 393-F	5'-GTGGTCTTCTACTTGTGTCAATAC
DYS 393-R	5'-AACTCAAGTCCAAAAAATGAGG
DYS 438-F	5'-TGGGGAATAGTTGAACGGTAA
DYS 438-R	5'-GTGGCAGACGCCTATAATCC

STR analysis by Genetic analyzer

A method for multiplex analysis of three STR loci (D13S631, D18S51 and D21S11) is described. The primers for the three loci are described in Fig. 9.5. One primer of each pair is labeled with a fluorescent dye. Since the amplified products of the three loci do not overlap the three primers are labeled with the same fluorescent dye (FAM).

(A) D13S317 (GenBank: G09017.1)

TGGGATGGGTTGCTGGACATGGTATC**ACAGAAGTCTGGGATGTGGA**GGAGAGTTCA
TTTCTTTAGTGGGCATCCGTGACTCTCTGGACTCTGACCCATCTAACGCCTATCTG
TATTTACAAATACAT**TATCTATCTATCTATCTATCTATCTATCTATCTATCT**
ATCTATCTATCAATCATCTATCTATCT**TTCTGTCTGTCTTTTTGGGC**TGCCTATGG
CTCAACCCAAGTTGAAGGAGGAGATTTGACCAACAATTCAAGCTCTCTGAATATGT
TTTGAA

(B) D18S51 (GenBank: AP001534.2)

GTTGAAAGATGAAATAACTTACTGAAATTGTTAATGAAGTATTGGATAAGCTACTTTAAAAA
TAA**CAAACCCGACTACCAGCAAC**AACACAAATAAACAAACCGTCAGCCTAAGGTGGACATGT
TGGCTTCTCTCTGTTCTTAACATGTTAAAATTAAAATTAACTTCTCTGGTGTGTGGAGATGT
CTTACAATAACAGTTGCTACTATTTCTTTTCTTTTTCTCTTTCTTTCCTCTCTCTTTTT**CTT**
TCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTC
TTTCTGAGACAAGGTCTCAATTTGTCACTCAGAGTGAAGTG**CAGTGGCATGAACATGGCTC**A
CTGCAGCCTTAACCTTCTGGGCTCAAGAACTCCTCCTGCCTCAGCCCTGCAAGTAGCTGAGA
CTACAGGCACGTGC

(C) D21S11 (GenBank: AP000433.2)

CTTGGACAGCCACACTGCCAGCTTCCCTGATTCTTCAGCTTGTAGATGGTCTGTTA
TGGGACTTTTCTCAGTCTCCATAAATATG**TGAGTCAATTCCCCAAGTGAA**TTGCCT
TCTATCTATCTATCTATCTGTCTGTCTGTCTGTCTGTCTATCTATCTATATC
TATCTATCTATCATCTATCTATCCATATCTATCTATCTATCTATCTATCTATCTAT
CTATCTATCTATCTATCGTCTATCTATCCAGTCTATCTACCTCCTATTA**GTCTGTC**
TCTGGAGAACATTGACTAATACAACATCTTTAATATATCACAGTTTAATTTCAAGT
TATATCATACCACTTCATACATTATATAAAACCTTACAGTGTTTCTCCCTTCTCAG

Fig. 9.5. PCR primers for D13S631, D18S51, D21S11 STR loci.

Loci:	D13S631, D18S51, D21S11	
Primer mix:	Make the following primer mixes:	
D13S317-F	6 FAM 5'-5'-ACAGAAGTCTGGGATGTGGA	10 pmol/µl
D13S317-R	5'-GCCCAAAAAGACAGACAGAA	10 pmol/µl
D18S51-F	6 FAM 5'-CAAACCCGACTACCAGCAAC	10 pmol/µl
D18S51-R	5'-GAGCCATGTTCATGCCACTG	10 pmol/µl
D21S11-F	6 FAM 5'-TGAGTCAATTCCCCAAG TGAA	10 pmol/µl
D21S11-R	5'-AGTCAATGTTCTCCAGAGACAGAC	10 pmol/µl
Amgl-F	6 FAM 5'-CTGATGGTTGGCCTCAAGCCT	10 pmol/µl
Amgl-R	5'- ATGAGGAAACCAGGGTTCCA	10 pmol/µl

For each sample label four tubes A, B, C and D (one each for the STR locus and Amgl).

Reaction conditions:

PCR mix:	23 µl
Respective primer mix:	1 µl
Taq polymerase:	0.1 µl (0.5 units)
DNA:	1 µl (20ng)

Thermal cycling:

Initial denaturation:	1 minute at 94°C
No. of cycles:	25
Denaturation:	48 seconds at 93°C
Annealing:	48 seconds at 60°C
Extension:	1 minute at 72°C
Final extension:	3 minute at 72°C

Pool 5 µl amplified product from each of the 4 reactions and use 1.3 µl of the pool.

HiDi formamide:	10 µl/sample
GeneScan Liz:	0.5 µl/sample
Denature in thermal cycler:	94°C for 4 minutes

Run on a Genetic analyzer as per manufacturers instructions.

Some technical problems with STR analysis

PCR stutter products

In amplification of STRs the elongating DNA strand can slip one or more steps forward resulting in formation of end products that are one or more repeat units smaller in length. This is also called slippage synthesis. On gel electrophoresis the shorter amplified products appears as shadow bands or peaks often called stutter bands or peaks. The stutter products larger in length than the actual product can also result by backward slippage but these are uncommon.

Stutter bands/peaks are most prominent in the amplification of di-nucleotide repeats but these are not uncommon in PCR of tetra-nucleotide repeats (Fig. 13.6).

Twin peaks of "+A" and "–A" products

Taq polymerase has the property of adding +A at the 3' end of the elongating strand that results in formation of products one base pair longer. These products appear as twin peaks on a genetic analyzer.

Quantity of DNA

Too much or too little DNA in PCR can produce poor results. The quantification of DNA is discussed in chapter 2. The quantity of DNA is very critical for analysis on a genetic analyzer. The analyzer works best with DNA quantity ranging from 0.5-1ng per reaction.

The gel electrophoresis and silver staining are more robust as far as the quantity of DNA is concerned. The usual range for good results is 50-200ng per reaction. The minimum amount of DNA that can be seen on a gel by this method is ~20ng (Fig. 9.6).

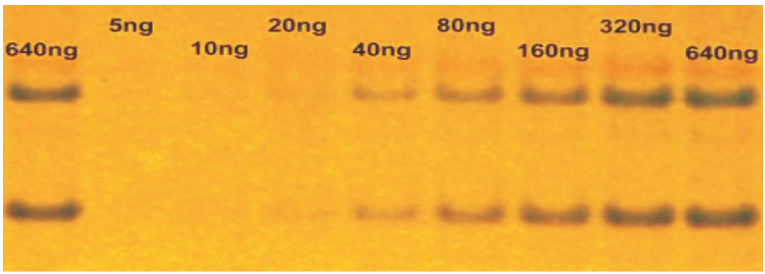


Fig. 9.6 Amplification of STR locus D8S1179 by using decreasing concentration of DNA. The 6% polyacrylamide gel was stained in silver nitrate. Minimum detectable amount of DNA is ~20ng.

Bibliography

1. Short Tandem Repeat DNA Internet DataBase, http://www.cstl.nist.gov/strbase/
2. Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B (1998) Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet 62: 1408-1415.
3. Smith RN (1995) Accurate size comparison of short tandem repeat alleles amplified by PCR. Biotechniques 18: 122-128.

10. Genomic Sequencing

The process of determining the sequence of an unknown DNA is called sequencing. There are many approaches for DNA sequencing. In the last couple of decades automated Sanger's di-deoxy chain termination method has dominated the scene. It has played a pivotal role in the completion of human genome project. More recently the new generation sequencing (NGS) technologies are emerging to produce enormous amount of sequencing data at very low cost. Up to one billion short reads of sequence can be produced in one instrument run.

Next generation sequencing (NGS)

The NGS technology involves template preparation, sequencing, imaging, and data analysis. The template preparation is done by randomly breaking genomic DNA into smaller pieces from which fragment or mate-pair templates are created by emulsion PCR (emPCR) or solid phase amplification. The templates are immobilized on a solid surface as spatially distinct sites on which thousands to billions of separate sequencing reactions can take place. The sequencing is done by dye (fluorescent) labeled nucleotides. These nucleotides are incorporated specifically and are cleaved during fluorescence imaging. DNA synthesis is terminated by reversible terminators. The imaging of fluorescent labeled nucleotides produces data that are aligned with a known reference sequence.

Sequencing by automated Sanger's method

In a diagnostic lab sequencing is mostly done to see the presence of any mutation or polymorphism in an already known sequence of DNA. The later is also called re-sequencing. Sequencing of an unknown DNA is a tedious process and its description is beyond the scope of this book.

During synthesis of DNA deoxy-nucleotides (dNTPs) are incorporated in the DNA strand. In Sanger's method DNA strand synthesis is terminated by incorporation of di-deoxy-nucleotides (ddNTPs). The ddNTPs are synthetic analogs of dNTPs that can bind at 5' end with the preceding dNTP but are unable to form bond at 3' end with the next dNTP. In DNA sequencing the reactions are terminated randomly resulting in generation of varying lengths of DNA strands each terminating at the respective ddNTP. Previously the sequencing was done by radio-labeled ddNTPs in four separate reaction vials and the end products were electrophoresed in four separate lanes. In the automated genetic analyzers ddNTPs are labeled with four different flurochromes and the reaction product is run in a single capillary tube.

In a typical sequencing reaction the target DNA is first amplified by a pair of primers. The amplified product is purified to remove excess dNTPs, and unused primers. The sequencing is initiated by a single “sequencing primer” which is complementary to the target region of interest. The sequencing primer may be one of the primers used in the initial PCR amplification of the fragment or any other primer complementary to region of interest in the amplified product. The reaction is done in a PCR tube that contains the amplified target DNA, the sequencing primer, dNTPs, fluorescent labeled ddNTPs, sequencing polymerase, and a reaction buffer. The end products of the reaction are purified and are analyzed on a genetic analyzer.

Direct sequencing

DNA can be sequenced directly by using single stranded or double stranded templates. Single stranded DNA is generated by asymmetric PCR in which the two primers are used in unequal proportions (one primer is in excess of the other). Double stranded DNA amplified by a pair of primers can be sequenced directly after separating its strands by heat denaturation and snap cooling.

Cycle sequencing

The DNA to be sequenced is first amplified by a pair of primers. The amplified target is then subjected to linear or asymmetric amplification by a single primer i.e. "sequencing primer". In a thermal cycler the target DNA and the reaction mixture are subjected to repeat cycling of denaturation, primer annealing, and extension. In each cycle DNA strands of varying lengths are produced that terminate at each ddNTP (Fig. 10.1). At the end of cycling the DNA strands of varying lengths are analyzed by capillary electrophoresis in a genetic analyzer.

(A)

3′-GAGTCGTACGCTACCATGTCATATCAACCATAGTTCCGATCCTAGTACTAAC-5′
5′-**CATGCGATGGTACAGTATAG**-3′→

3′-GAGTCGTACGCTACCATGTCATATCAACCATAGTTCCGATCCTAGTACTAAC-5′
5′-**CATGCGATGGTACAGTATAG**-**ddTTP**
5′-**CATGCGATGGTACAGTATAG**T-**ddTTP**
5′-**CATGCGATGGTACAGTATAG**TT-**ddGTP**
5′-**CATGCGATGGTACAGTATAG**TTG-**ddGTP**
5′-**CATGCGATGGTACAGTATAG**TTGG-**ddTTP**
5′-**CATGCGATGGTACAGTATAG**TTGGT-**ddATP**
5′-**CATGCGATGGTACAGTATAG**TTGGTA-**ddTTP**
5′-**CATGCGATGGTACAGTATAG**TTGGTAT-**ddCTP**
5′-**CATGCGATGGTACAGTATAG**TTGGTATC-**ddATP**
5′-**CATGCGATGGTACAGTATAG**TTGGTATCA-**ddATP**
5′-**CATGCGATGGTACAGTATAG**TTGGTATCAA-**ddGTP**
5′-**CATGCGATGGTACAGTATAG**TTGGTATCAAG-**ddGTP**

(B)

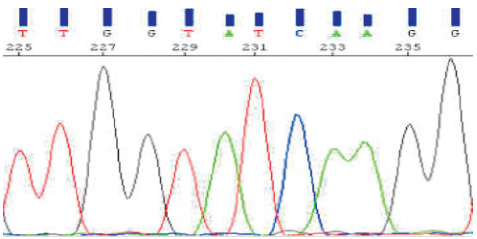


Fig. 10.1. (A) DNA template and the sequencing primer. The newly formed DNA strand is terminated when a ddNTP instead of the dNTP is incorporated. The four ddNTPs are tagged with different fluorescent dyes. (B) The DNA strands of varying lengths are separated by capillary electrophoresis. Each fragment is represented by a coloured peak on electropherogram shown in the lower half of the picture.

Example

Locus: β-globin gene
GenBank accession: NG_000007.3

PCR Amplification:

- Forward primer 5'-ACCTCACCCTGTGGAGCCAC
- Reverse primer 5'-CCCCTTCCTATGACATGAACTTAA
- Amplified product: 676bp
- Reaction volume 50 µl
- Primer concentration: 2 µl each of the forward and the reverse primer (10 pmol/µl)
- Taq polymerase: 1 unit (0.2µl)
- Template DNA: 2 µl (~500ng)
- Thermal cycling:
  - Initial denaturation: 1 minute at 94°C
  - Cycles: 25
  - Denaturation: 1 minute at 94°C
  - Annealing: 1 minute at 65°C
  - Extension: 1 minute 30 seconds at 72°C
  - Final extension: 3 minute at 72°C
- Electrophoresis: To see the quality of amplification run 3µl of amplified product on 6% mini-polyacrylamide gel for 40 minutes at 150 volts.
- Staining: 0.1% silver nitrate

Purification of amplified product

The amplified product contains unused dNTPs and primers that must be removed before sequencing. This may be done by gel purification, enzyme digestion or column purification.

1. Gel purification

a. It is tedious but efficient in removing non-specific amplified products.
b. The amplified DNA is run on 2% agarose gel at 150 volts for one hour. The gel is stained in ethidium bromide and is visualized by UV light. The bands of amplified DNA are localized and are carefully cut with knife. The piece of gel is transferred to an Eppendorf tube. The DNA may be extracted from the gel by a commercial spin column.

2. Enzymatic digestion

a. The amplified DNA is incubated in a single step with exonuclease and alkaline phosphatase. However, the

method is unable to remove non-specific amplified products.

b. In a PCR tube 2-5µl of amplified product is incubated with 1µl each of exonuclease and alkaline phosphatase at 37°C for 15 minutes. The enzymes are inactivated at 80°C for another 15 minutes. The product is ready for use.

3. Column purification

a. Micro-columns with silica gel filters are efficient in removing unwanted contaminants in amplified DNA. These may also be used to concentrate DNA. However, these are unable to remove non-specific amplified products.

b. The micro-columns usually take 100µl of the amplified product. The initial amplification is done in duplicate of 50µl each and the end products are pooled.

c. Purification is done as per instruction of the commercial column manufacturer.

Cycle sequencing

1. Since the cycle sequencing reaction requires extensive optimization the procedure is best done by a commercial sequencing kit.
2. The choice of sequencing primer depends on the application. Usual choice is the "forward primer" used in the initial amplification. However, any primer complementary to the amplified target sequence can be used. The sequencing primer is used at a concentration recommended by the sequencing kit manufacturer.
3. Volume of the purified amplified product (template) for use in cycle sequencing reaction depends on the quality and quantity of initial amplification. It ranges from 2µl for very good amplification to 8µl for poor amplification. For better results DNA in the template may be quantified (Chapter 2).
4. Thermal cycling parameters are also set as recommended by the kit manufacturer.

Purification of sequencing reaction products

The cycle sequencing reaction product contains many unwanted substances like unused dNTPs and fluorescent labeled ddNTPs. The later may cause dye blobs to appear in the electropherograms. These unwanted substances must be completely removed before the sample is run on a genetic analyzer.
Commercial kits are available for clean up of the sequencing reaction products. A good low cost method based on ethanol precipitation is described below.

Ethanol precipitation method

1. Add the following to the tube containing sequencing reaction product:
   a. Absolute ethanol, 50µl
   b. EDTA 125mM, 2µl
   c. Sodium acetate 3M, 2µl
2. Vortex gently to mix and centrifuge briefly.
3. Incubate at room temperature for 15 minutes.
4. Centrifuge at 14000 rpm for 15-25 minutes.
5. Carefully remove the supernatant using a fine tipped pipette.
6. Add 60µl 70% ethanol.
7. Vortex briefly and centrifuge at 14000 rpm for 5 minutes.
8. Carefully remove the supernatant using a fine tipped pipette.
9. Add 10µl Hi-Di formamide and use for electrophoresis.

Run conditions for genetic analyzer

The run conditions are chosen as per manufacturer's instructions.

Bibliography

1. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-termination inhibitors. Proc Natl Acad Sci, USA 74: 5463-5467.
2. Murray V (1989) Improved double-stranded DNA sequencing using the linear polymerase chain reaction. Nucleic Acids Res 17: 8889.
3. Metzker ML (2010) Sequencing technologies-the next generation. Nature Rev Genet 11: 31-46.

11. Quality Control in Diagnostic PCR

In a PCR lab the quality control is mainly concerned with the control of errors in performance of tests and verification of test results. It includes internal and external quality control. The internal quality control includes monitoring of laboratory performance by using control material and/or repeat measurements.

Errors in sample collection, transport and nucleic acid extraction

DNA is fairly stable when it is in the cell. The cellular DNA may get degraded by DNAse derived from bacterial contamination of the sample. The sample for DNA extraction should be processed before bacteria grow in it. The sample could be collected in a sterile container. But if the sample is already contaminated then it's collection in s sterile container would not help. DNA is a very large and fragile molecule and it can also be broken by vigorous shaking during extraction. The degraded DNA is not good for PCR and this is an important reason for false negative or poor quality of results.

RNA is far more sensitive to degradation than DNA. It can easily be degraded by RNAse present in the environment especially those derived from the cells from where RNA is being extracted. Since RNA isolation procedure takes place in a strong denaturant that renders RNAse inactive, the RNA integrity is mostly at risk prior to and after the isolation. Therefore in RNA extraction the handling of sample prior to isolation and the storage of isolated RNA are very critical.

Inhibitors of PCR

Heparin, porphyrin, SDS, phenol, and proteinase-K are potent inhibitors of PCR. SDS and phenol must be completely removed from the sample in DNA extraction. The blood samples should be collected in EDTA. Heparin should not be used as anticoagulant as it inhibits PCR. Proteinase K can be inactivated by heating. PCR may also be inhibited by chocolate that might contaminate the DNA extracted from a mouthwash. The effect of inhibitors can be reduced or abolished by diluting the sample as it also dilutes the inhibitor.

Errors in PCR

There are numerous factors that may affect the efficiency of PCR. The quality of reagents, primers, buffer pH, quality and quantity of DNA and PCR tubes etc. are all important.

1. The quality of Taq polymerase is important in diagnostic work. The cheaper brands of Taq may be alright for research work but in diagnostic work these could give false positive results.
2. The primer, especially its design, is the most important determinant in specificity of PCR (Chapter 3).
3. The concentration of DNA in PCR is also important. Apart from the quality, too much or too little DNA can give false negative results.
4. While setting up PCR it is always better to premix reagents. If ten samples are to be amplified prepare reaction mix in one master tube and then transfer to individual tubes labeled 1-10. This helps in avoiding inconsistency in pipetting etc. The DNA should always be

added at the end.

5. The quality of PCR tubes is often ignored. The tube should be of very thin wall that allows quick transfer of heat. The tube should fit well in the sample block. A loose fitting tube would have air around it that could interfere with heat transfer.

False positive PCR results

PCR is an extremely sensitive technique. In thirty cycles one molecule of DNA can be amplified a billion fold. The amplified product provides an ideal structure for re-amplification. It can be carried over from one sample to another and may act as target for further amplification. The carry over is a problem that is serious as well as difficult to handle.

Identifying false positives

The false positives are identified by including good quality negative, positive and non target controls (reagent blank). In diagnostic PCR it is mandatory to include at least one negative and a positive DNA control. It is also essential to include a reagent blank that contains every thing except DNA (non target control). The latter is included to exclude contamination of the reagents by extraneous DNA. If the reagent blank shows amplification of the target DNA the result of complete batch of samples becomes null and void.

Preventing false positives

1. Sample preparation, amplification and end point analysis should be physically separated. The samples may be prepared in a separate room or in a bio-safety cabinet.
2. The place should be irradiated with UV light form time to time or when contamination is suspected. UV light can destroy DNA rendering it unsuitable for further amplification. Clean lab benches and equipment periodically with freshly diluted 10% bleach solution.
3. Separate sets of pipettes may be used for the three steps. Handling of DNA before amplification is not as serious a cause of carry over from one tube to another. It is the amplified DNA that usually causes more serious problems. In a resource constrained set-up if separate pipettes for every step can not be used then at least the amplified DNA should be handled with separate pipettes clearly marked “amplified DNA only”. Barrier filter tips are also useful in preventing the carry over due to aerosol spray of the amplified products.
4. Disposable gloves should be worn to prevent carryover of amplified DNA to the sample preparation area.
5. Re-amplification of any contaminating amplified DNA from a previous PCR can be avoided if dUTP instead of dTTP is used in PCR. The dUTP containing amplified products can be rendered un-amplifiable by enzymatic degradation of uracil with UNG (Uracil DNA Glycosylase).

Handling false positives

One should take all possible precautions in preventing contamination. Once developed it is difficult to handle. To investigate errors due to

contamination by extraneous DNA proceed as follows:

1. The reagents most often contaminated include those used in extraction, working solutions of primes, PCR buffer and Taq polymerase. Less often the stock solutions of PCR buffer, dNTPs and primers etc. may also be contaminated.
2. In a stepwise manner carry out the PCR with "in-use" reagents but replacing one item at a time from the fresh stock.
3. Discard the contaminated reagent if, by luck, one is able to find it!
4. In a busy diagnostic lab there may not be enough time to investigate the source of contamination. In that case the best choice would be to discard all remaining quantities of the working reagents and prepare fresh from the stocks. Use fresh vial of Taq polymerase. In the mean while de-contaminate the work area with UV light for 30-60 minutes. The pipettes are the usual source of carry over and should be thoroughly irradiated with UV light.

False negatives

The false negatives are also serious problems that often do not get enough attention. These are best detected by inclusion of good quality positive, negative and internal controls. Inclusion of an internal PCR control is always good to reduce the risk of false negatives. This is usually done by including a separate pair of primers in the reaction tube to amplify an unrelated part of the DNA. A PCR result that shows positive internal control but no amplification of the target DNA is called a true negative.

The false negatives may be caused by any of the following:

1. Degraded, too little or too much DNA.
2. Poor quality of reagents.
3. Failure to add any constituent of reaction e.g. primer, Taq polymerase or DNA.
4. Problem with the PCR machine.

Avoiding false negatives

1. Observe precautions in extraction of DNA/RNA.
2. Use only the prescribed amount of DNA/RNA.
3. In diagnostic work only good quality of reagents should be used that are not expired. In resource constrained labs reagents may be used beyond expiry if their potency is proven by previous runs or in positive controls.
4. The reagents like PCR buffers and primers etc. should be stored in aliquots. This helps in reducing loss of quality due to repeat freeze thawing. Moreover if an aliquot of reagents is contaminated it can be conveniently discarded.
5. Wearing disposable gloves helps in preventing DNAses present on the skin from degrading DNA.
6. The PCR mix contains several ingredients whose quality and quantity can be detrimental. For example insufficient amount of Mg^{++} can result in poor amplification. Increasing the concentration of Mg^{++} can increase amplification but too much of Mg^{++} can result in false positives. Similarly poor quality of dNTPs, primers, and Taq polymerase can also cause false negatives.

Instrument calibration

Instrument calibration gets least attention in a PCR lab. Thermal cyclers like any other lab equipment require calibration from time to time. The commonest error is the difference in the temperature of the block/reaction tube and that displayed on the screen. If the block temperature is lower than the displayed temperature primers can anneal non-specifically. The result would be non-specific amplification. Higher block temperature than the displayed would result in higher denaturation and annealing temperatures. Former could rapidly denature Taq polymerase while the latter would result in difficulties in primer annealing. The net result is reduction in the amplification.

The thermal cyclers should be periodically checked for the block and the displayed temperatures. Good quality digital temperature meters should be kept in a PCR lab. Use the following procedure to check the temperature in the reaction tube:

1. Place five PCR tubes in the sample wells of the heating block (four tubes in each corner well and the fifth tube in the centre well).
2. Add 50µl water to each of the tubes and keep their lids open.
3. Turn on the thermal cycler and set to hold temperature for five minutes at 65°C.
4. Once the temperature has reached 65°C allow one minute for temperature of water in the tubes to reach 65°C.
5. One by one measure the temperature in each of the five tubes by dipping the sensor in to the water in each tube.
6. The temperature in all of the tubes should be equal and at 65°C.
7. The machine needs calibration if there is any difference of temperature between the tubes and that displayed on the screen. Calibration is done by entering the new calibration factor in the software if this right is given to the user. Otherwise an engineer's help would be required. If none of the options are available one could use temperature settings in the thermal cycling programme making "+" or "-" adjustment for the difference in temperatures of the tubes and the display.
8. If there is any difference in temperature of individual tubes the machine needs servicing by a qualified engineer.

Bibliography

1. Kwok S, Higuchi R (1989) Avoiding false positives with PCR. Nature 339: 237-238.
2. Rys PN, Persing DH (1993) Preventing false positives: Quantitative evaluation of three protocols for inactivation of polymerase chain reaction amplification products. J Clin Microbiol 31: 2356-2360.

12. Setting-up a PCR Lab

Since the applications of PCR range very widely it is advisable to establish separate sections for genetic, malignant and infectious disorders. The forensic DNA lab in any case needs to be established separately. PCR is a relatively new technique for most diagnostic pathology laboratories. Therefore more than often one is required to create a space for PCR facility in an already established lab. No matter how small the PCR setup would be it should be housed in a separate room. The size of the lab depends on the range of applications and the quantum of workload.

Whether the PCR lab is small, medium or large or is aimed to cater for genetic, malignant, or infectious disorders the basic format remains the same. Considering the extreme sensitivity of PCR it is essential to divide or at least restrict the flow of work to three separate areas. These include areas for specimen preparation, PCR setup and post PCR processing. A suggested layout for a medium sized PCR lab is shown in Fig 12.1.

Specimen preparation area

The specimen preparation area is dedicated for receiving and processing all specimens. Special precautions are also required if any of the infectious samples are being received. No handling of amplified product should be done in this area. The equipment used in this area especially the pipettes etc. should not be used in other areas of the lab.

PCR setup area

This is an area where PCR is setup. It should be maintained free of contamination from sample preparation and post PCR processing areas. The pipettes etc. should not be used at other areas. If possible each worker should have its own bench space and set of pipettes etc. This area is also used for reagent preparation and storage.

Post PCR processing area

This area is used for PCR amplification and processing of the amplified products. The thermal cyclers may also be placed in this area. This is potentially the most dangerous area that may spread contamination to the rest of the work areas. The pipettes used in this area must be marked “amplified DNA only”.

In a small lab the entire procedure of PCR can be done in a single room. The areas may be demarcated or physically separated by creating partitions. Whether PCR is done in separate rooms or separate areas in a single room the flow of work should remain uni-directional.

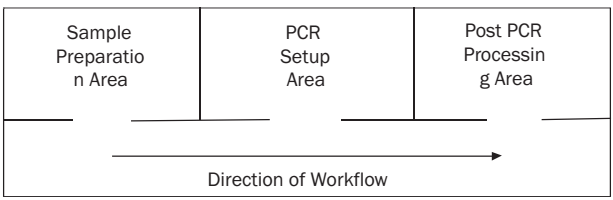


Fig 12.1 Suggested layout of a PCR laboratory.

Equipment required for a PCR lab

The basic equipment required for a PCR lab is listed in Table 12.1.

Table 12.1 List of basic equipment for a medium sized PCR lab.

Equipment	Quantity
DNA Thermal cycler	2
Micro-centrifuge (up to 15000 rpm)	2
Vortex mixer	2
Analytical balance (0.000g)	1
Adjustable pipettes (full range)	4 sets
Poly-acrylamide gel electrophoresis apparatus	2
Submarine agarose gel electrophoresis tanks	2
DC Power Supply (100-300 volts)	2
Water bath	1
Gel dryer with vacuum pump	1
Gel documentation system	1
Refrigerator (-20°C)	2
Real time PCR machine (optional)	1

Thermal cycler

It is the most important piece of equipment in a PCR lab. A wide variety of instruments are available with widely ranging prices. In general machines that can accommodate greater numbers of tubes and have higher ramping rates (ability to rapidly cool or heat) are more expensive. Such machines are usually suitable for labs with very large quantum of work.

One should buy a machine that best suits the requirement. If the daily work load of a lab is between 10-20 samples there is no point in buying an expensive machine with capacity to hold 96 samples or to have very high ramping rates. In a diagnostic lab a spare PCR machine should be available as backup.

Real time PCR machine

A fairly good number of manufacturers are making real time PCR machines. A machine with capacity to detect larger number of fluorescent dyes (colours) is more expensive. While choosing a real time PCR machine one should keep the requirements in mind. A diagnostic lab that is mostly involved in real time PCR of infectious disorders should buy a machine with two colour detection capability. It is good enough for doing most TaqMan® probe based real time PCR applications.

Measures to protect against power breakdowns

1. Frequent power breakdowns can adversely affect the overall functioning of the lab. All sensitive equipment especially the thermal cyclers must have a backup electric supply. Un-interrupted Power Supply (UPS) that generates square wave electric current can be

harmful for the equipment with moving parts. All such equipment should have a sine wave UPS. The UPS should be of appropriate power. Lab equipment like incubators, water baths, and water distillation plants that use maximum electric power should not be placed on UPS.

2. Loose power connections can be a source of instrument malfunction. Adequate attention must be paid to electric wiring. Loose connections and inappropriate electric wiring can also be a fire hazard in a lab.
3. In a large or a medium sized lab a central generator supply with UPSs for individual instruments can be a good combination to tackle power breakdowns. Refrigerators used for storage of reagents/kits should also be put on generator supply. The generator used should have the capacity well in excess of the total power load. It may be advisable to keep the high power consuming equipment like air conditioners and water baths etc. off the generator supply.

Section-II

13. PCR Applications in Inherited Disorders

The PCR technology has enabled even small labs to develop expertise in the diagnosis of common inherited disorders. The most useful application of PCR in such situations is the prenatal diagnosis (PND). There is a long list of inherited disorders where PCR has a significant role to play. A typical example is β-thalassaemia. It is common and its molecular genetics is also well known. Trisomies, although not inherited in most situations, are other major areas where PCR can play useful role in rapid prenatal diagnosis.

In this chapter molecular diagnosis of β-thalassaemia is described as an example. The diagnosis of other single gene disorders caused by point mutations e.g. cystic fibrosis etc. can also be done on similar lines.

β-thalassaemia

There are over 200 different mutations that cause β-thalassaemia. However, each ethnic population has its own set of common and uncommon mutations. The first step in carrying out PCR based diagnosis of β-thalassaemia is to know the pattern of mutations in the target population. This knowledge is essential for genetic diagnosis.

Screening for β-thalassaemia mutations:

Most of the β-thalassaemia mutations are point mutations that are best detected by ARMS (Chapter 7). ARMS can be done in separate reactions for each mutation or as multiplex PCR.

Screening strategy

Standard ARMS (separate reactions)

1. In the first round of PCR common mutations are tested.
2. The samples that do not show any of the common mutations are then tested in the next rounds for the uncommon and the rare mutations.
3. The list and the sequences of primers for the β-thalassaemia mutations in the major world populations are given in Table 13.1. For convenience the primers may be given a serial number. The numbers are easier to remember and document than the full nomenclature of the allele.
4. In each reaction the ARMS primer for a mutation is run with another primer called “common primer”.
5. Most of the ARMS primers are designed as complementary to the forward strand of DNA. All of these primers are used with the same “common primer” which is complementary to the reverse strand (primer 3).

Table 13.1. ARMS primers for β-thalassaemia mutations in the major world populations. The sequence of primers for the mutant and the normal alleles are given. Each ARMS primer is used with one of the common primers. (No. 3 & 4). The amplified product size is also given. Primers 1 and 2 are used to amplify 861bp PCR internal control fragment of the distal portion of β-globin gene.

Control and common primers			
1. Control-F	5'-CAATGTATCATGCCTCTTTGCACC		
2. Control-R	5'-GAGTCAAGGCTGAGAGATGCAGGA		
3. Common-1	5'-ACCTCACCCTGTGGAGCCA		
4. Common-2	5'-CCCCTTCCTATGACATGAACTTAA		
Allele	**Primer sequence**	**Used with**	**Product Size**
Asian mutations			
Fr 8-9 (+G) M	5'-CCTTGCCCCACAGGGCAGTAACGGCACACC	3	215
Fr 8-9 N	5'-CCTTGCCCCACAGGGCAGTAACGGCACACT	3	215
IVSI-5 (G-C) M	5'-CTCCTTAAACCTGTCTTGTAACCTTGTTAG	3	285
IVSI-5 N	5'-CTCCTTAAACCTGTCTTGTAACCTTGTTAC	3	285
Fr 41-42 (TTCT) M	5'-GAGTGGACAGATCCCCAAAGGACTCAACCT	3	439
Fr 41-42 N	5'-GAGTGGACAGATCCCCAAAGGACTCAAAGA	3	443
Del 619bp-F	5'-CAATGTATCATGCCTCTTTGCACC	2	242
Del 619bp-R	5'-GAGTCAAGGCTGAGAGATGCAGGA	1	242
Cd 15 (GA) M	5'-TGAGGAGAAGTCTGCCGTTACTGCCCAGTA	4	500
Cd 15 N	5'-TGAGGAGAAGTCTGCCGTTACTGCCCAGTG	4	500
Cd 5 (-CT) M	5'-ACAGGGCAGTAACGGCAGACTTCTCCGCGA	3	205
Cd 5 N	5'-ACAGGGCAGTAACGGCAGACTTCTCCGCAG	3	205
IVSI-1 (G-T) M	5'-TTAAACCTGTCTTGTAACCTTGATACGAAA	3	281
IVSI-1 N	5'-GATGAAGTTGGTGGTGAGGCCCTGGGTAGG	4	450
Cd 30 (G-C) M	5'-TAAACCTGTCTTGTAACCTTGATACCTACG	3	280
Cd 30 (G-A) M	5'-TAAACCTGTCTTGTAACCTTGATACCTACT	3	280
Cd30 N	5'-TAAACCTGTCTTGTAACCTTGATACCTACC	3	280
Fr 16 (-C) M	5'-TCACCACCAACTTCATCCACGTTCACGTTC	3	238
Fr 16 N	5'-TCACCACCAACTTCATCCACGTTCACGTTG	3	239

6. ARMS primers for some mutations are complementary to the reverse strand and these are used with another "common primer" complementary to the forward strand (primer 4).
7. In each reaction two additional primers (1 & 2) are used to amplify 861bp fragment of the distal end of β-globin gene. This works as an internal control in the ARMS PCR. The location of the primers 1 & 2 is such that it also allows identification of 619bp deletion in the β-globin gene. If the del 619bp mutation is present the control fragment, instead of the usual 861bp, is reduced to 242bp.

Multiplex ARMS

1. A multiplex ARMS PCR for 12 β-thalassaemia mutations found in the Pakistani population is described. Similar multiplexes for the other ethnic populations can also be developed.
2. Three combinations of primers for multiplex ARMS of β-thalassaemia are shown in Table 13.2. The multiplexes AD-1 and AD-2 contain the ARMS primers used with "common primer 3" while the AD-3 contains the ARMS primers that are used with "common primer 4".
3. The amplified products of all the mutations tested in each multiplex are sufficiently different for resolution by mini polyacrylamide gel electrophoresis. However, the size of the fragments generated by IVSI-1 and IVSI-5 differ by only 5bp and are difficult to resolve on a mini gel. Similarly there is no difference between the fragments of Cd30 (G-C), Cd30 (G-A) and IVSI-1 mutations. The problem of differentiating IVSI-1 and IVSI-5 is overcome by adding IVSI-1 primer to AD-1 and AD-2 multiplexes. IVSI-5 results in amplification with AD-1, but IVSI-1 causes amplification with AD-1 and AD-2. Amplification with AD-2 but not AD-1 indicates Cd30. The difference between Cd30 (G-C) and Cd30 (G-A) is only of academic interest because the same normal primer is used to differentiate between the homozygotes and heterozygotes of the two mutations. The AD-3 primer combination includes Cd15 and Cap+1 and both of these primers are used with "common primer 4".
4. An allelic ladder for the respective mutations is prepared by pooling the PCR products of separately amplified reactions of various mutations. The allelic ladder is kept frozen in aliquots. 5µl of the pooled product is used in all polyacrylamide gel electrophoresis runs.
5. The amplified products of the multiplex ARMS and the respective allelic ladders are run on 6% polyacrylamide mini gels.
6. Interpretation of the multiplex ARMS results is shown in Fig. 13.1.

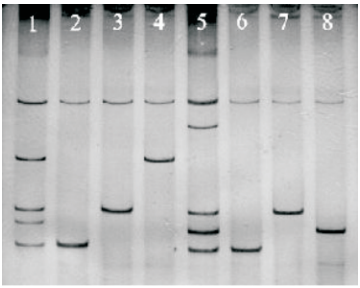


Fig. 13.1. Silver stained mini polyacrylamide gel electrophoresis of multiplex ARMS PCR products. Lane 1 and 5 show allelic ladders for the AD-1 and AD-2 multiplexes. All lanes show the 861bp internal control fragment. The sample in lane 2, 3 and 4 are positive for Fr 8-9 (+G), IVSI-5 (G-C) and Fr 41-42 (-TTCT) mutations respectively. Lanes 6, 7 and 8 show Cd5 (-CT), IVSI-1 (G-T) and Fr 16 (-C) mutations respectively.

Table 13.2. Multiplex ARMS primer combinations for β-thalassaemia mutations found in the Pakistani population.

Primer ID:	Mutations:	Product Size:
AD -1	Fr 8-9(+G)	215bp
	IVSI -5(G-C)	285bp
	Fr 41-42(-TTCT)	439bp
	IVSI -1(G-T)	280bp
	619bp	242bp
AD -2	Cd 5(-CT)	205bp
	Fr 16(-C)	238bp
	IVSI -1(G-T)	280bp
	Cd 30(G-C)	280bp
	Cd 30(G-A)	280bp
	IVSII -1(G-A)	634bp
AD -3	Cd 15(G-A)	500bp
	Cap+1(A-C)	567bp

Homozygous or heterozygous mutation

1. Once the mutation is identified the next step is to find whether the mutation is homozygous or heterozygous. This is done by setting up a separate PCR reaction in which the normal allele of the respective mutation is tested. The list and the sequences of the normal ARMS primers are shown in Table 13.1.
2. In a homozygote either two copies of the same mutation (true homozygote) or two different mutations (compound heterozygote) are seen.
3. The true homozygote shows only one mutation. Its homozygosity is ascertained by testing for the normal allele checked in a separate PCR reaction. The compound heterozygote shows two different mutations.
4. A heterozygote has the mutant as well as the normal allele. There is no need to test for the normal allele if the individual is known to have thalassaemia trait.

Sequencing of β-globin gene

The β-globin gene sequencing is done to identify an unknown mutation or to confirm the results of ARMS PCR. The gene is approximately 1.5 kb in length. On the 36cm capillary of genetic analyzer approximately 600 bases can be sequenced in one go. The entire β-globin gene can be sequenced in overlapping segments.

Sequencing strategy

Most of the β-thalasaemia mutations are located in the first two exons and the intervening sequence (segment-I). The segment-II mostly covers the IVS-II which is an unusual site for the mutations. The segment-III covers the third exon and the terminal portion of the gene. It also contains only a few uncommon mutations.

- Segment-I
- Target region: Promoter region & Exon-I
    - Forward primer: 5'-AGACATAATTTATTAGCATGCATG
    - Reverse primer: 5'-CTCCTTAAACCTGTCTTGTAACCT
    - Fragment size: 809 bp
    - Sequencing primer: 5'-AGACATAATTTATTAGCATGCATG
- Segment-II
    - Target region: Exon-I, IVS-I & Exon-II
    - Forward primer: 5'-ACCTCACCCTGTGGAGCCAC
    - Reverse primer: 5'-CCCCTTCCTATGACATGAACTTAA
    - Fragment size: 676 bp
    - Sequencing primer: 5'-ACCTCACCCTGTGGAGCCAC
- Segment-III
    - Target region: Exon-III & Poly-A site
    - Forward primer: 5'-CAATGTATCATGCCTCTTTGCACC
    - Reverse primer: 5'-GAGTCAAGGCTGAGAGATGCAGGA
    - Amplified fragment: 861 bp
    - Sequencing primer: 5'-CAATGTATCATGCCTCTTTGCACC

The details of the sequencing protocol are given in Chapter 10.

Results:

The results of normal sequence, IVSI-5 (G-C) and Fr 8-9 (+G) mutation are shown in Fig 13.2 and 13.3.

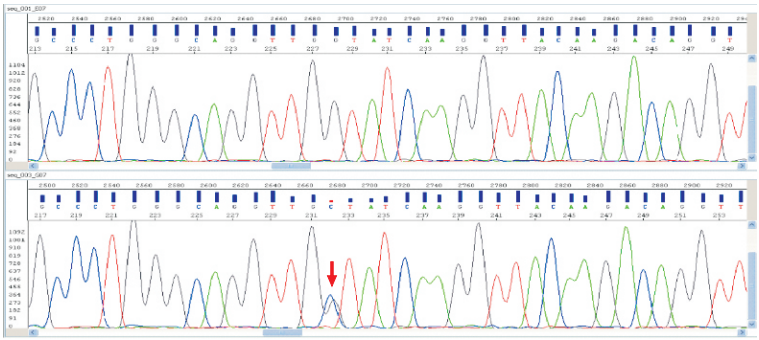


Fig 13.2. Electropherogram of the β-globin gene showing the sequence around the first exon-intron junction. The upper half shows the normal sequence whereas IVSI-5 (G-C) substitution (arrow) can be seen as overlapping peaks (black and blue) in the lower part of the picture.

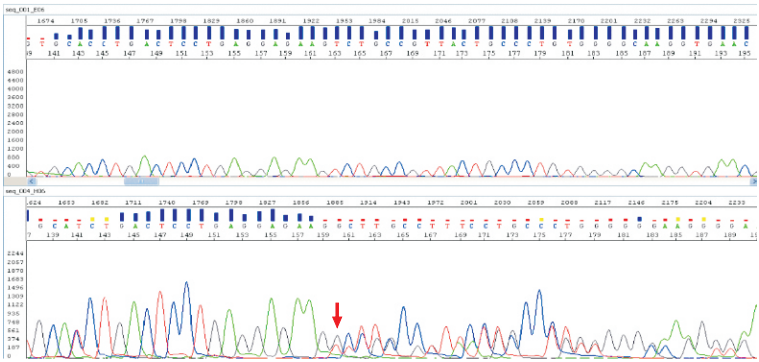


Fig 13.3. Upper half of the picture shows normal sequence of β-globin gene whereas Fr 8-9 (+G) insertion is seen to cause complete disruption of the sequence beyond codon eight (arrow) in the lower part of the picture.

Prenatal Diagnosis (PND)

PND by direct mutation analysis

1. Identify the parent's mutations by testing the DNA of each parent or an affected child if any of the parents is not available.
2. Carefully dissect the chorionic villus sample (CVS) to remove any maternal decidua.
3. Extract DNA from the CVS (fetus) (Chapter 2).
4. Setup ARMS PCR as follows (also see chapter 7):

1. If both parents have the same mutation label ten tubes and add:
   - i. Tubes1-5: primers 1+2 + common primer + mutation primer.
   - ii. Tubes 6-10: primers 1+2 + common primer + normal primer for the mutation.

iii. In tubes 1-5 add DNA of father, mother, CVS (in duplicate) and negative control for the mutation.
iv. In tubes 6-10 add DNA of CVS (in duplicate), positive control for normal allele (normal DNA), negative control for normal allele (homozygous for the mutation) and no DNA (reagent blank).
v. Gel electrophoresis and the interpretation of the results are shown in Fig. 13.4.

2. If the two parents have different mutations label nine tubes and add:
   i. Tubes1-4: primers 1+2 + common primer + primer of father's mutation
   ii. Tubes 5-9: primers 1+2 + common primer + primer of mother's mutation
   iii. In tubes 1-4 add DNA of father, CVS (in duplicate) and negative control for father's mutation.
   iv. In tubes 5-9 add DNA of mother, CVS (in duplicate), negative control for mother's mutation and no DNA (reagent blank).

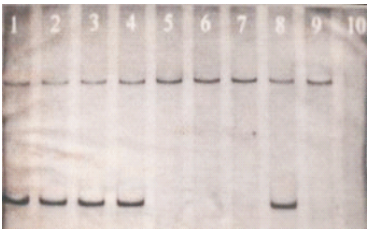


Fig. 13.4. Polyacrylamide gel electrophoresis of prenatal diagnosis for beta thalassaemia. All lanes except lane 10 show 861bp internal control bands. Absence of any result in the reagent blank (Lane 10) excludes the false positive results. Lanes 1 & 2 shows the parent's mutations (IVSI-5). Lanes 3 & 4 show the same mutation (IVSI-5) tested in duplicate in the fetal DNA. Lane 5 is a negative control for IVSI-5 mutation. Lanes 6 & 7 shows testing for the normal allele of IVSI-5. Its absence in the fetal DNA confirms that the fetus has homozygous IVSI-5. Lanes 7 and 8 are negative and positive controls for the normal allele of IVSI-5 respectively. Lane 8 is a negative control for the IVSI-5 allele.

Exclusion of maternal contamination

1. Presence of maternal tissue in the CVS can be a potential source of error in prenatal diagnosis. Since the mother is a carrier who has the mutant as well as the normal alleles, contamination of CVS by the maternal tissue would make all normal or homozygous fetal results to appear as thalassaemia trait.
2. When the parents have the same mutation and the fetal diagnosis is "normal" or "homozygous (thalassaemia major)" maternal contamination is ruled out.
3. When the parents have two different mutations and the mother's mutation is not present in the CVS, maternal contamination is ruled out.
4. When the parents have the same mutation and the fetal diagnosis is "heterozygous (thalassaemia trait)" maternal contamination can not be ruled out.

5. Experience has shown that meticulous cleaning of the CVS is enough to safeguard against any errors due to maternal contamination.
6. The maternal contamination in the CVS can be tested by short tandem repeat (STR) analysis.
7. The CVS and the maternal DNA are run for various STR loci (Chapter 9 & 15).
8. The STR marker is called "informative" when its alleles can distinguish between the maternal and the fetal DNA. For example at D21S11 locus if the mother has alleles 28,29 and the fetus (CVS) has 28,31. The marker is informative because allele 29 and 31 are exclusive for the mother and the fetus respectively. It also indicates that there is no maternal contamination. It would have indicated maternal contamination had the exclusive maternal allele (29) also been seen in the fetus which is usually expressed as 28,31(29).
9. STR locus D21S11 is often used because of its high degree of polymorphism. It's limitation is that if the fetus has trisomy 21 and the extra chromosome 21 is of maternal origin the result could appear as maternal contamination.
10. If the SRT is not informative at any locus the analysis is done at other loci.

PND of a genetic disorder by linkage analysis

In couples when the parent's mutation is not identified PND can also be done by linkage analysis. There are several single nucleotide polymorphisms (SNP) and short tandem repeats (STR) that are closely linked to several important genetic disorders. The polymorphic loci are inherited en-block with the gene of interest and can be used to track their inheritance.

Linkage analysis by SNP

The SNPs are recognized by the strategic use of restriction enzymes. In the first step DNA of the parents and the affected child are tested to establish linkage of the SNP with the chromosome carrying the mutation. Once an informative marker is identified the fetal DNA is tested to determine its genotype. By convention the SNP is written as "+" when it is present and as "-" when it is not present (Chapter 8).

Pre-requisites of linkage based prenatal diagnosis:

1. The couple must have a previously affected child.
2. At least one fully informative marker should be available. The marker is called informative when both of the parents are heterozygous (-/+).
3. Linkage based prenatal diagnosis is not possible if any of the parents is homozygous (+/+) for the marker.
4. Linkage based prenatal diagnosis may be possible by using more than one partially informative markers (when only one of the parents is heterozygous for the marker).
5. The linkage based diagnosis may turn out to be incorrect if during the

meiotic cross over marker and the gene of interest dissociate (separate) from each other.

Example:

β-globin gene complex contains at least seven SNPs that can be used to carry out linkage based prenatal diagnosis of thalassaemia. One of the SNPs can be recognized by restriction enzyme Hinc-II. The results of genotyping at this SNP in a couple with an affected child are explained in Fig. 13.5.

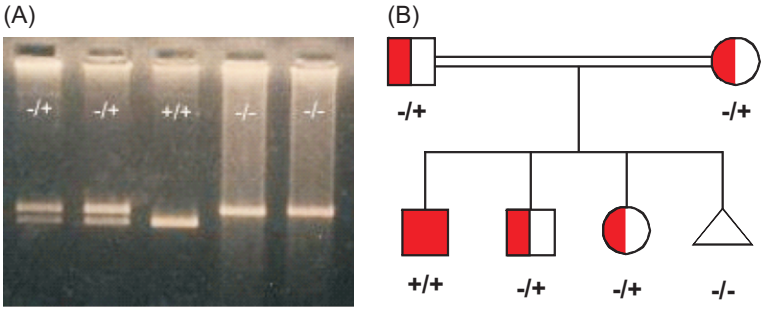


Fig. 13.5. (A) Ethidium bromide stained agarose gel electrophoresis of Hinc-II digested fragments of β-globin gene. Father and the mother have -/+ genotype. The affected child of the couple has +/+ genotype indicating that the chromosome carrying the + site in the father and the mother has the β-thalassaemia mutation. The fetal DNA has -/- genotype indicating not affected by β-thalassaemia.

Linkage analysis by STR

Linkage analysis may also be done by STR markers. Several important STR loci are known to be present in many important genes. Duchene muscular dystrophy (DMD) is a good example in which at least four di-nucleotide STRs present in the Introns 44, 45, 49 and 50 of the dystrophin gene on X-chromosome can be used. The intra-genic STRs have an additional advantage as their failure to amplify indicates deletions and inversions in the dystrophin gene (Fig. 13.6).

Example

Primers for amplification of di-nucleotide STRs in the four introns of dystrophin gene.

DMD-44 F	5'-TCCAACATTGGAAATCACATTTCAA
DMD-44 R	5'-TCATCACAAATAGATGTTTCACAG
DMD-45 F	5'-GAGGCTATAATTCTTTAACTTTGGC
DMD-45 R	5'-CTCTTTCCCTCTTTATTCATGTTAC
DMD-49 F	5'-CGTTTACCAGCTCAAAATCTCAAC
DMD-49 R	5'-CATATGATACGATTCGTGTTTTGC
DMD-50 F	5'-AAGGTTCCTCCAGTAACAGATTTGG
DMD-50 R	5'-TATGCTACATAGTATGTCCTCAGAC

Each pair of primer for the respective intron is used for amplification in a separate tube as follows:

- Reaction volume: 25 µl
- Primer concentration: 1 µl (5 pmol each/µl)
- Taq polymerase: 0.5 units (0.1µl)
- Template DNA: 1µl (200ng)
- Thermal cycling:
  - Initial denaturation: 1minute at 94°C
  - No. of cycles: 25
  - Denaturation: 30 seconds at 94°C
  - Annealing: 30 seconds at 62°C
  - Extension: 2 minutes at 65°C
  - Final extension: 3 minute at 72°C
- Electrophoresis: 16 X 20 X 0.1cm 6% polyacrylamide gel. Load 3µl amplified product in 3µl loading dye Run at 100 volts overnight.
- Staining: 0.1% Silver nitrate.
- Result: Fig 13.6

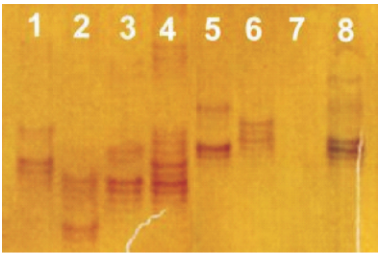


Fig. 13. 6. PAGE after PCR amplification of di-nucleotide repeats in Intron 44 (Lanes 1-4) and Intron 45 (Lanes 5-8) of dystrophin gene in a family with DMD. Lanes 1 & 5 show father's sample, lanes 2 & 6 show mother's sample, lanes 3 & 7 show sample of the child affected by DMD, and lanes 4 & 8 show sample of fetal DNA. Sample of the affected child in lane 3 shows amplification product whereas the same sample in lane 7 does not show any amplification indicating a deletion in the dystrophin gene involving intron 45. The fetal DNA sample (lanes 4 & 8) does not show the deletion in intron 45. The result of Intron 44 can also be used for linkage analysis. The father (lane 1) has one allele due to one X chromosome (a faint band of PCR stutter product can also be seen). The mother (lane 2) shows two alleles due to two X chromosomes. The affected child (lane 3) being a male has one allele that is inherited from the mother. This indicates that out of the mother's two alleles the larger allele corresponding to that of the affected child can be used as a marker for the abnormal X chromosome. The fetal sample (lane 4) shows two alleles one inherited from the father and the other from the mother. The result is consistent with female fetus carrier for DMD.

Limitations of linkage based PND

1. The couple should have a living affected child.
2. At least one informative marker must be present.
3. There is approximately 1% chance that during meiotic cross-over the marker and the abnormal gene could dissociate causing error in diagnosis.
4. Presence of mutation out-side the abnormal gene can cause error in diagnosis.

Sources of error in PND

1. Maternal contamination in fetal sample
2. PCR failure
3. Clerical mistakes
4. Meiotic crossover in linkage analysis
5. Non paternity

Screening for Trisomies

Chromosomal trisomies can be detected by STR analysis. Trisomies 13, 18 and 21 constitute over 90% of the clinically significant problems. The three trisomies can be quickly screened by PCR amplification of STR loci. The principle of screening is that each chromosome carries one allele of an STR. In a trisomy three instead of the usual two alleles are seen. A technical problem often arises if two or all three of the alleles on the three chromosomes have the same number of repeat units. The problem is solved either by testing additional loci or by densitometry or automated STR analysis on genetic analyzer (Chapter 9).

The following allele patterns may be seen at an STR locus (example D21S11):

1. Normal

   a. Disomic di-allelic (28,29)
   b. Disomic mono-allelic (28,28)

2. Abnormal

   a. Trisomic tri-allelic (28,29,30)
   b. Trisomic di-allelic (28,29,29)
   c. Trisomic monoallelic (28,28,28)

The STR allele sizing is done by PCR amplification of the locus followed by gel electrophoresis. In manual allele sizing gel electrophoresis is done with an allelic ladder. The results may be interpreted by visual inspection of the gel or by densitometry. The trisomy screening can be done without allele sizing if the amplified DNA samples of the father, mother and the subject are run simultaneously on a polyacrylamide gel (Fig. 13.7) The DNA amplified with a fluorescent labeled primer when run on genetic analyzer also provides information on the allele sizes (Fig. 13.8).

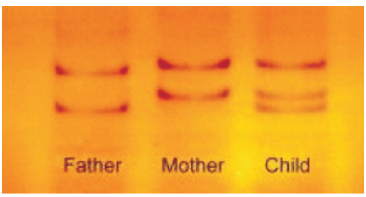


Fig. 13.7. Polyacrylamide gel electrophoresis of PCR amplification at D21S11 locus in a couple suspected to have child affected by trisomy 21. The father and the mother have a disomic diallelic pattern where as the child shows trisomic triallelic pattern confirming trisomy 21. In addition it is also visible that the child inherited the extra chromosome 21 from the mother.

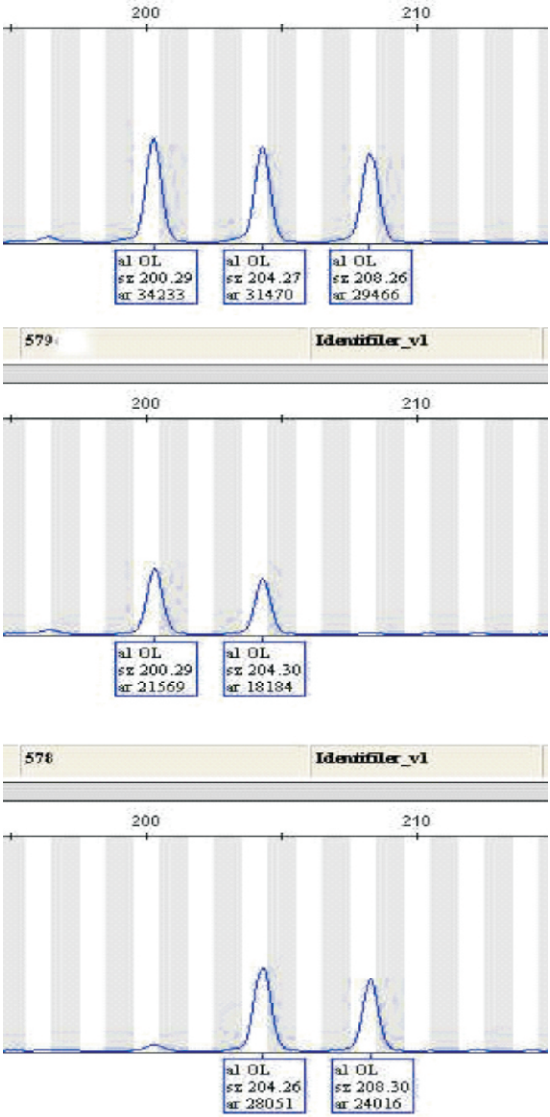


Fig. 13.8. PCR amplification at D21S11 locus ran on a genetic analyzer. The father and the mother have a disomic diallelic pattern whereas the affected child (top) shows trisomic triallelic pattern confirming trisomy 21.

Primers used for screening of trisomies:

Trisomy 13:

D13S317-F	5'-ACAGAAGTCTGGGATGTGGA
D13S317-R	5'-GCCCAAAAAGACAGACAGAA
D13S634-F	5'-TCCAGATAGGCAGATGATTCAAT
D13S634-R	5'-CCTTCTTCTTCCCATTGATA

Trisomy 18

D18S51-F	5'-CAAACCCGACTACCAGCAAC
D18S51-R	5'-GAGCCATGTTCATGCCACTG
D18S535-F	5'-TCATGTGACAAAAGCCACAC
D18S535-R	5'-AGACAGAAATATAGATGAGAATGCA

Trisomy 21

D21S11-F	5'-GTGAGTCAATTCCCCAAG
D21S11-R	5'-GTTGTATTAGTCAATGTTCTCC
D21S1411-F	5'-ATGATGAATGCATAGATGGATG
D21S1411-R	5'-AATGTGTGTCCTTCCAGGC
D21S1412-F	5'-CGGAGGTTGCAGTGAGTT
D21S1412-R	5'-GGGAAGGCTATGGAGGAGA
D21S1414-F	5'-AAATTAGTGTCTGGCACCCAGTA
D21S1414-R	'-CAATTCCCCAAGTGAATTGCCTTC

Limitations

Detection of trisomies by PCR is only a screening method and it is not a substitute for cytogenetic analysis. Its major limitation is that in PCR the smaller alleles may be preferentially amplified giving the false impression of trisomy.

Bibliography

1. Old JM, Varawalla NY, Weatherall DJ (1990) Rapid detection and prenatal diagnosis of -thalassaemia: studies in Indian and Cypriot populations in the UK. Lancet 336: 834-837.
2. Clemens PR, Fenwick RG, Chamberlain JS, Gibbs RA, Andrade M, Chakraborty R (1991) Carrier detection and prenatal diagnosis in Duchenne and Becker muscular dystrophy families, using dinucleotide repeat polymorphisms. Am J Hum Genet 49: 951-960.
3. Ahmed S, Petrou M, Saleem M (1996) Molecular genetics of -thalassaemia in Pakistan: a basis for prenatal diagnosis. Br J Haematol 94: 476-481.
4. Old JM (1996) Haemoglobinopathies. Prenat Diagn 16: 1181-1186.
5. Adinolfi M, Pertl B, Sherlock J (1997) Rapid detection of aneuploidies by microsatellites and the quantitative fluorescent polymerase chain reaction. Prenat Diagn 17: 1299-1311.
6. Huisman TJH, Carver MFH (1998) The δ- and β-thalassaemia repository. Haemoglobin 22: 169-195.
7. Ahmed S, Saleem M, Petrou M, Sultana N, Raashid Y, Waqar A, Anwar M, Karamat KA (2000) Prenatal diagnosis of -thalassaemia in Pakistan: experience in a Muslim country. Prenat Diagn 20: 378-383.

14. PCR Applications in Neoplastic Disorders

At the sub-cellular level cancer is certainly a genetic disorder. The molecular genetics of cancer is typically marked by mutation(s) in the oncogenes, tumour suppressor genes or micro RNA genes. These genetic changes carry potential for diagnosis of cancer by PCR. PCR can be used to diagnose the disease, predict its prognosis or monitor the disease progress and response to treatment if a cancer specific DNA marker is identified. PCR can also be used to demonstrate the association of some malignancies and viruses e.g. human papillomavirus and the cervical cancer and HTLV-1 infection and leukaemia.

Ig gene rearrangement

Immunoglobulin and T-cell receptor genes in the germ-line configuration are made of several individual parts called variable (V), diversity (D) and joining (J) segments. On antigenic challenge the germ line segments rearrange to form a much smaller gene. The un-rearranged Ig gene is too large for amplification by PCR. It could be amplified if the gene rearrangement brings the PCR primers for the target DNA very close to each other.

The lymphoid malignancies show clonal rearrangements of Immunoglobulin genes or T-cell receptor genes. The assembly of IgH genes from the variable (VH), diversity (DH), and joining (JH) regions creates a DNA target that is complex and unique to each patient. In the subsequent paragraphs a simple broad spectrum PCR assay is described that can pick up to 90% of the IgH gene rearrangements in patients of lymphoproliferative disorders. A forward primer complementary to the consensus JH sequence is used with a reverse primer in the VH region.

Forward primer (JH) (5pmol):	5'-AACTGCAGAGGAGACGGTGACC
Reverse primer (VH) (5pmol):	5'-CTGTCGACACGGCCGTGTATTACT
Reaction volume:	25µl
PCR buffer:	22µl
Taq polymerase:	0.5 units (0.1µl)
Primer mix:	1µl (5pmol of each primer/µl)
DNA:	2µl (~200ng)

Thermal cycling

- Initial denaturation 94°C for 5 minutes.
- Thirty cycles each of:
  - i. Denaturation at 94°C for 1 minute
  - ii. Annealing at 60°C for 1 minute
  - iii. Extension at 72°C for 1 minute 30 seconds
- Final extension at 72°C for 3 minutes

Electrophoresis may be done on 2% agarose at 150 volts for 60 minutes or 6% mini-polyacrylamide gel at 150 volts for 30 minutes (Fig. 14.1).

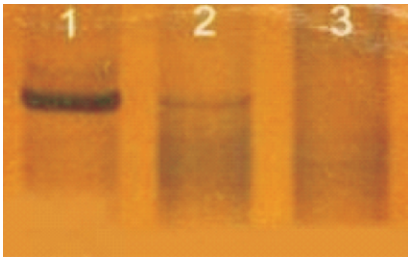


Fig. 14.1. PCR for Immunoglobulin heavy chain gene rearrangement. Lane 1 shows a sharp band of clonal Ig gene rearrangement where as lane 3 shows a diffuse pattern of polyclonal gene rearrangements. Lane 2 shows a faint clonal band of minimal residual disease along with a diffuse polyclonal background in the post treatment sample of the patient at lane 1.

Real time PCR for IgH gene rearrangement

Forward primer (JH) (5pmol):	5'-AACTGCAGAGGAGACGGTGACC
Reverse primer (VH) (5pmol):	5'-CTGTCGACACGGCCGTGTATTACT
SYBR green ready reaction mix:	23µl
DNA:	2µl (~200ng)

SYBR green ready reaction mix containing dNTPs and Taq polymerase is available from many commercial sources.

Thermal cycling

- Initial denaturation 94°C for 5 minutes.
- Thirty cycles each of:
  - i. Denaturation at 94°C for 1 minute
  - ii. Annealing at 60°C for 1 minute (read fluorescence)
  - iii. Extension at 72°C for 1 minute 30 seconds

Melting curve analysis may be added at the end of the real time PCR protocol to differentiate between the specific and the non-specific amplification products.

Bcl-II gene rearrangement

Follicular lymphoma is characterized by translocation of bcl-II gene from chromosome 18 to IgH gene on chromosome 14. The fusion gene is a lymphoma specific marker that can be used in diagnosis and monitoring of disease activity.

mbr	5'-GAGTTGTACGTGGCCTG
mcr	5'-CGCTTGACTCCTTTACGTGC
s-icr	5'-TCGTTCTCAGTAAGTGAGAGTGC
LJH	5'-TGAGGAGACGGTGACC

Real Time PCR

Primers:

Mbr-F	5'-TTAGAGAGTTGCTTTACGTGGCC
IgH-R	5'-ACTCACCTGAGGAGACGGTGAC
Mbr-Probe	6 FAM 5'-TTTCAACACAGACCCACCCAGAGCC-TAMRA
Primer mix:	Make the primer mix of the four primers and two probes at concentration of 5pmol each (Chapter 3).

PCR protocol:

Reaction volume:	25µl
PCR mix:	22µl
Primer mix:	1µl
Taq polymerase:	0.5 units (0.1µl)
DNA:	2µl

Thermal cycling

- Initial denaturation 95°C for 10 minutes
- Forty cycles each of:
  - i. Denaturation at 95°C for 15 seconds
  - ii. Annealing/extension at 60°C for 1 minute
- Read fluorescence at 60°C step

Bcr-abl gene rearrangements

A reciprocal translocation between the long arms of chromosomes 9 and 22 (Philadelphia chromosome) is found in >90% patients of chronic myeloid leukemia and 15–25% patients with acute lymphoblastic leukemia. This translocation transposes c-abl oncogene from chromosome 9q34 to bcr gene on chromosome 22q11. The fused bcr/abl gene provides specific marker for diagnosis and disease monitoring. The translocation can be detected by first making cDNA of the bcr-abl mRNA followed by PCR.

Common bcr-abl gene rearrangements

Over 95% of the bcr-abl transcripts in CML are between introns 13/14 on bcr gene and intron 2 on abl gene ($e_{13}a_2$ & $e_{14}a_2$). The transcripts result in protein called p210. In ALL 70% of the transcripts are e1a2 that result in protein called p190. About 25% of transcripts in ALL are $e_{13}a_2$ or $e_{14}a_2$ (Fig. 14.1). In a minority of the patients atypical transcripts may also be observed.

PCR screening strategy

Sequence of the bcr-abl gene transcripts and the PCR primers with TaqMan® probe are shown in Fig. 14.2. Three separate forward primers E-1, E-13 and E-14 are used with a common reverse primer A-2. A common bcr-abl TaqMan® probe complementary to the abl gene is used. A fragment of the abl gene just close to the breakpoint region is amplified as an internal control by a separate pair of primers.

A. Homo sapiens partial mRNA for bcr-abl e14a2 chimeric protein (GenBank: AJ131466.1) .

ACTGTCCACAGCATTCCGCTGACCATCAATAAGGAAGATGATGAGTCTCCGGGGCTCTATGGGTTTCTG
AATGTCATCGTCCACTCAGCCACTGGATTTAAGCAGAGTTCAAAAGCCCTTCAGCGGCCAGTAGCATCT
GACTTTGAGCCTCAGGGTCTGAGTGAAGCCGCTCGTTGGAACTCCAAGGAAAACCTTCTCGCTGGACCC
AGTGAAAATGACCCCAACCTTTTCGTTGCACTGTATGATTTTGTGGCCAGTGGAGATAACACTCTAAGC
ATAACTAAAGGTGAAAAGCTCCGGGTCTTAGGCTATAATCACAATGGGGAATGGTGTGAAGCCCAAACC
AAAAATGGCCAAGGCTGGGTCCCAAGCAACTACATCACGCCAGTCAACAGTCTGGAGAAACACTCCTGG
TACCATGGGCCTGTGTCCCGCAATGCCGCTGAGTATCTGCTGAGCAGCGGGATCAATGGCAGCTTCTTG
GTGCGTGAGAGTGAGAGCAGTCCTGGCCAGAGGTCCATCTCGCTGAGATACGAAGGGAGGGTGTACCAT

B. Homo sapiens partial mRNA for bcr-abl e13a2 chimeric protein (GenBank: AJ131467.1).

AAGAAGTGTTTCAGAAGCTTCTCCCTGACATCCGTGGAGCTGCAGATGCTGACCAACTCGTGTGTGAAA
CTCCAGACTGTCCACAGCATTCCGCTGACCATCAATAAGGAAGAAGCCCTTCAGCGGCCAGTAGCATCT
GACTTTGAGCCTCAGGGTCTGAGTGAAGCCGCTCGTTGGAACTCCAAGGAAAACCTTCTCGCTGGACCC
AGTGAAAATGACCCCAACCTTTTCGTTGCACTGTATGATTTTGTGGCCAGTGGAGATAACACTCTAAGC
ATAACTAAAGGTGAAAAGCTCCGGGTCTTAGGCTATAATCACAATGGGGAATGGTGTGAAGCCCAAACC
AAAAATGGCCAAGGCTGGGTCCCAAGCAACTACATCACGCCAGTCAACAGTCTGGAGAAACACTCCTGG
TACCATGGGCCTGTGTCCCGCAATGCCGCTGAGTATCTGCTGAGCAGCGGGATCAATGGCAGCTTCTTG
GTGCGTGAGAGTGAGAGCAGTCCTGGCCAGAGGTCCATCTCGCTGAGATACGAAGGGAGGGTGTACCAT

C. Homo sapiens bcr-abl e1a2 chimeric protein mRNA, partial cds (GenBank: AF113911.1).

GACTTCTCCTCTGGCCAGTCCAGCCGCGTGTCCCCAAGCCCCACCACCTACCGCATGTTCCGGGACAAA
AGCCGCTCTCCCTCGCAGAACTCGCAACAGTCCTTCGACAGCAGCAGTCCCCCCACGCCGCAGTGCCAT
AAGCGGCACCGGCACTGCCCGGTTGTCGTGTCCGAGGCCACCATCGTGGGCGTCCGCAAGACCGGGCAG
ATCTGGCCCAACGATGGCGAGGGCGCCTTCCATGGAGACGCAGAAGCCCTTCAGCGGCCAGTAGCATCT
GACTTTGAGCCTCAGGGTCTGAGTGAAGCCGCTCGTTGGAACTCCAAGGAAAACCTTCTCGCTGGACCC
AGTGAAAATGACCCCAACCTTTTCGTTGCACTGTATGATTTTGTGGCCAGTGGAGATAACACTCTAAGC
ATAACTAAAGGTGAAAAGCTCCGGGTCTTAGGCTATAATCACAATGGGGAATGGTGTGAAGCCCAAACC
AAAAATGGCCAAGGCTGGGTCCCAAGCAACTACATCACGCCAGTCAACAGTCTGGAGAAACACTCCTGG
TACCATGGGCCTGTGTCCCGCAATGCCGCTGAGTATCTGCTGAGCAGCGGGATCAATGGCAGCTTCTTG
GTGCGTGAGAGTGAGAGCAGTCCTGGCCAGAGGTCCATCTCGCTGAGATACGAAGGGAGGGTGTACCAT

Fig. 14.2. Homo sapiens partial mRNAs for bcr-abl chimeric proteins (A-C). The PCR primers and the TaqMan probes for the bcr-abl gene rearrangement and a control fragment of the abl gene are shown. The bcr gene sequences are shown in blue whereas the abl gene sequences are shown in red.

RNA extraction

For RNA extraction commercial TRIzol® Reagent® (Invitrogen, USA) as described in Chapter 2 gives good quality results.

Preperation of cDNA

Primer ABL-R is used for preparing the cDNA. Oligo dT primer for poly-A tail of mRNA or random hexamers may also be used for this purpose. In this experiment RNA extracted from a patient known to have bcr-abl gene rearrangement was serially diluted ranging from 100% to 50%, 25%, 12.5%, 6.25%, 3.12%, and 1.56%. A known negative RNA was also included as a negative control. cDNA was prepared from each dilution and the negative control as per the following protocol:

Primer (ABL-R) 5'-GGCCACAAAATCATACAGTGCA

Reactions conditions (per reaction):

- 5 X reaction buffer: 4 µl
- 10 mmol dNTPs mix: 2 µl
- RT Primer (10 pmol/µl): 1 µl
- MMLV- Reverse Transcriptase (200 U/µl): 1 µl
- RNAse inhibitor (20 U/µl): 1 µl
- Deionized water: 11 µl
- RNA: 2 µl
- Incubation: 42°C for 60 minutes
- RT inactivation: 70°C for 5 minutes

Real Time PCR

Primers and Probes:

BCR-ABL	
E-1	5'-GCAGATCTGGCCCAACGAT
E-13	5'-GCATTCCGCTGACCATCAATAA
E-14	5'-CAGCCACTGGATTTAAGCAGAGT
A-2 (R)	5'-TCCAACGAGCGGATTCACT
BCR-ABL-P	6 FAM 5'-AAGCCCTTCAGCGGCCAGTAGCATCT-BHQ1

Internal control

ABL-F	5'-GCTGGGTCCCAAGCAACTAC
ABL-R	5'-ACACAGGCCCATGGTACCA
ABL-P	JOE 5'-TCACGCCAGTCAACAGTCTGGAGAAACA-BHQ1

Primer mix: Make the primer mix of the six primers and the two probes at concentration of 5pmol each (Chapter 3).

PCR protocol (per reaction):

Reaction volume:	25µl
PCR mix:	21µl
Primer mix:	1µl
Taq polymerase:	0.5 units (0.1µl)
cDNA:	3µl

Thermal cycling

i. Initial denaturation 95oC for 10 minutes
j. Forty cycles each of:
    i. Denaturation at 95°C for 15 seconds
    ii. Annealing/extension at 60°C for 1 minute
k. Take fluorescence reading at 60°C step

Result

The real time plot of the serially diluted sample and the negative control is shown in Fig. 14.3.

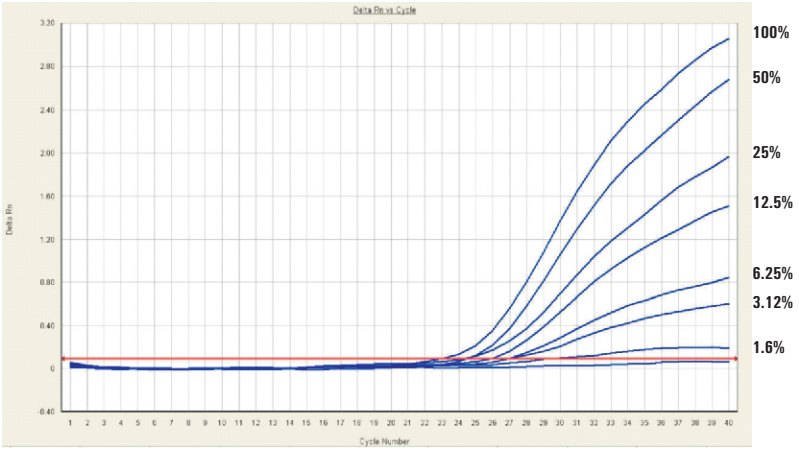


Fig. 14.3. Real time plot of the serially diluted RNA sample (100-1.56%) positive for the bcr-abl gene rearrangement and a known negative sample.

A strategy for comprehensive molecular diagnosis of leukaemia

Molecular diagnosis of leukaemia may be helpful in prognostic stratification and detection of minimal residual disease. The WHO classification of acute leukaemia is based on demonstration of a large number of gene rearrangements that also correspond to gross chromosomal abnormalities. A comprehensive strategy for detection of various molecular genetic abnormalities in Acute Lymphoblastic Leukaemia (ALL) and Acute Myeloid Leukaemia (AML) is described.

RNA extraction

Commercial TRIzol® Reagent (Invitrogen, USA) as described in Chapter 2.

Preperation of cDNA

Primer: Random hexamers at 0.2 µg/µL

Reactions conditions (per reaction):

- 5 X reaction buffer: 4 µl
- 10 mmol dNTPs mix: 2 µl
- RT Primer (10 pmol/µl): 1 µl
- MMLV- Reverse Transcriptase (200 U/µl): 1 µl
- RNAse inhibitor (20 U/µl): 1 µl
- Deionized water: 11 µl
- RNA: 2 µl
- Incubation: 42°C for 60 minutes
- RT inactivation: 70°C for 5 minutes

Real Time PCR

Primers and Probes

ALL

E2A-PBX1	t(1;19)	(GenBank: M31222 & M86546)
Forward:	ENF-E2A	5'-CCAGCCTCATGCACAACCA
Reverse:	ENR-PBX1	5'-GGGCTCCTCGGATACTCAAAA
Probe:	ENP E2A	FAM-5'-CCCTCCCTGACCTGTCTCGGCC-BHQ1
MLL-AF4	t(4;11)	(GenBank: L04284 & L13773)
Forward1:	ENF-MLL	5'-CCCAAGTATCCCTGTAAAACAAAAA
Forward2:	ENF-MLL	5'-GATGGAGTCCACAGGATCAGAGT
Reverse:	ENR-AF4	5'-GAAAGGAAACTTGGATGGCTCA
Probe:	ENP-AF4	FAM-5'-CATGGCCGCCTCCTTTGACAGC-BHQ1
TEL-AML1	t(12;21)	(GenBank: U11732 & D43969)
Forward:	ENF-TEL	5'-CTCTGTCTCCCCGCCTGAA
Reverse:	ENR-AML1	5'-CGGCTCGTGCTGGCAT
Probe:	ENP-TEL	FAM-5'-TCCCAATGGGCATGGCGTGC-BHQ1
m-bcr	t(9;22) p190	GenBank: X02596 & X16416)
Forward:	ENF-BCR	5'-CTGGCCCAACGATGGCGA
Reverse:	ENR-ABL	5'-CACTCAGACCCTGAGGCTCAA
Probe:	ENP-ABL	FAM-5'-CCCTTCAGCGGCCAGTAGCATCTGA - BHQ1
M-bcr	t(9;22)	p210 (GenBank: X02596 & X16416)
Forward:	ENF-BCR	5'-TCCGCTGACCATCAAYAAGGA
Reverse:	ENR-ABL	5'-CACTCAGACCCTGAGGCTCAA
Probe:	ENP-ABL	FAM-5'-CCCTTCAGCGGCCAGTAGCATCTGA-BHQ1

AML

SIL-TAL1	del(1)	(GenBank: M74558 & S53245)
Forward:	ENF-SIL	5'-CGCTCCTACCCTGCAAACA
Reverse:	ENR-TAL	5'-CCGAGGAAGAGGATGCACA
Probe:	ENP-SIL	FAM-5'-ACCTCAGCTCCGCGGAAGTTGC-BHQ1
PML-RARA	t(15;17)	(GenBank: M73778 & X06538)
Forward1:	ENF-PML	5'-TCTTCCTGCCCAACAGCAA
Forward2:	ENF-PML	5'-ACCTGGATGGACCGCCTAG
Forward3:	ENF-PML	5'-CCGATGGCTTCGACGAGTT

Reverse:	ENR-RARA	5'- GCTTGTAGATGCGGGGTAGAG
Probe:	ENP-RARA	FAM-5'-AGTGCCCAGCCCTCCCTCGC-BHQ1
CBFB-MYH1	inv(16)	(GenBank: L20298 & D10667)
Forward:	ENF-CBFB	5'-CATTAGCACAACAGGCCTTTGA
Reverse1:	ENR-MYH11	5'-AGGGCCCGCTTGGACTT
Reverse2:	ENR-MYH11	5'-CCTCGTTAAGCATCCCTGTGA
Reverse3:	ENR-MYH11	5'-CTCTTTCTCCAGCGTCTGCTTAT
Probe:	ENP-CBFB	FAM-5'-TCGCGTGTCCTTCTCCGAGCCT-BHQ1
AML1-ETO	t(8;21)	(GenBank: D43969 & D14289)
Forward:	ENF-AML1	5'-CACCTACCACAGAGCCATCAAA
Reverse:	ENR-ETO	5'-ATCCACAGGTGAGTCTGGCATT
Probe:	ENP-AML1	FAM-5'-AACCTCGAAATCGTACTGAGAAGCACTCCA-BHQ1

Primer and Probe mixes:

Make the respective primer and probe mixes of each molecular defect in a separate tube at concentration of 5pmol each (Chapter 3).

PCR protocol (per reaction):

Reaction volume:	25µl
PCR mix:	21µl
Primer mix:	1µl
Taq polymerase:	0.5 units (0.1µl)
cDNA:	3µl

Thermal cycling

- Initial denaturation 95°C for 10 minutes
- Forty cycles each of:
    i. Denaturation at 95°C for 15 seconds
    ii. Annealing/extension at 60°C for 1 minute

Take fluorescence reading at 60°C step

P53 gene mutations:

The gene for tumor suppressor protein p53 is located on chromosome 17. It is mutated and/or deleted in >50% of human cancers. A large number of p53 gene mutations and deletions have been described in various cancers.

The mutations in p53 gene may be screened by DGGE, SSCP or melting curve analysis (Chapter 5). Final confirmation of the mutations is done by genomic sequencing (Chapter 10).

A strategy for screening of mutations in exon 5-8 of p53 gene by melting curve analysis is described:

Primers	Sequence and annealing conditions
Exon5a-F	5'-CAACTCTGTCTCCTTCCTCTTCCTAC
Exon5a-R	5'-AGCCATGGCACGGACGCG
	65–60 °C touchdown 0.5 °C/cycle for 10 cycles
Exon5b-F	5'-CTCCTGCCCGGCACCCGC
Exon5b-R	5'-CTAAGAGCAATCAGTGAGGAATCAGA
	65–60 °C touchdown 0.5 °C/cycle for 10 cycles
Exon6-F	5'-CAACCACCCTTAACCCCTCCT
Exon6-R	5'-AGACGACAGGGCTGGTTGC
	68–58 °C touchdown 1.0 °C/cycle for 10 cycles
Exon7-F	5'-AGGCGCACTGGCCTCATC
Exon7-R	5'-GAGGCTGGGGCACAGCA
	68–58 °C touchdown 1.0 °C/cycle for 10 cycles
Exon8-F	5'-GACCTGATTTCCTTACTGCCTCTTG
Exon8-R	5'-AATCTGAGGCATAACTGCACCCTT
	63.5–58.5 °C touchdown 0.5 °C/cycle for 10 cycles

For screening of mutations in each exon use the respective pair of primers in separate tubes.

Forward primer concentration:	5pmol
Reverse primer concentration:	5pmol
SYBR green ready reaction mix:	23µl
DNA:	2µl (~200ng)

Thermal cycling

- Initial denaturation 94°C for 5 minutes.
- Thirty cycles each of:
  - i. Denaturation at 94°C for 1 minute
  - ii. Annealing at 60°C for 1 minute (read fluorescence)
  - iii. Extension at 72°C for 1 minute 30 seconds
- Select melting curve analysis form software of real time thermal cycler

BRCA1 and BRCA2 mutations

BRCA1 & BRCA2 (BReast CAncer) tumour suppressor genes are located on chromosomes 17 and 13 respectively. A large number of mutations have been described in the two genes that are strongly associated with familial breast and ovarian cancer.

The mutations may be screened by DGGE, SSCP and melting curve analysis (Chapter 5). Final confirmation of the mutations is done by genomic sequencing (Chapter 10).

A strategy for screening of mutations in exon 2 & 20 of BRCA1 and exon 11 of BRCA2 genes by melting curve analysis is described:

Primers

BRCA1-F	5'-AAAAGATATAGATGTATGTTTTGCTAATGTGT
Exon 2-R	5'-TCCCAAATTAATACACTCTTGTGCTGA
BRCA1-F	5'-GAGTGGTGGGGTGAGATTTTTGTC
Exon 20-R	5'-CCTGATGGGTTGTGTTTGGTTTCT
BRCA2-F	5'-CGAAAATTATGGCAGGTTGTTACG
Exon 11-R	5'-GCTTTCCACTTGCTGTACTAAATCCA

For screening of mutations in each exon use the respective pair of primers in separate tubes.

Forward primer concentration:	5pmol
Reverse primer concentration:	5pmol
SYBR green ready reaction mix:	23µl
DNA:	2µl (~200ng)

Thermal cycling

- Initial denaturation 94°C for 5 minutes.
- Thirty cycles each of:
  - i. Denaturation at 94°C for 1 minute
  - ii. Annealing at 60°C for 1 minute (read fluorescence)
  - iii. Extension at 72°C for 1 minute 30 seconds
- Select melting curve analysis form software of real time thermal cycler

JAK-2 mutation (V617F)

A simple PCR for identification of JAK-2 mutation by ARMS is described:

Primers:

JAK2-Mutant-F (5pmol)	5'-AGCATTTGGTTTTAAATTATGGAGTATATT	
JAK2-Control-F (5pmol)	5'-ATCTATAGTCATGCTGAAAGTAGGAGAAAAG	
JAK2-Common-R (10pmol)	5'-CTGAATAGTCCTACAGTGTTTTCAGTTTCA	
Amplified products:	Mutation:	203bp
	Control:	364bp

Primer mix: Make the primer mix of the three primers at concentration of 5pmol each (Chapter 3).

Reaction volume:	25µl
PCR buffer:	22µl
Taq polymerase:	0.5 units (0.1µl)
Primer mix:	1µl (5pmol of each primer/µl)
DNA:	2µl (~200ng)

Thermal cycling

- Initial denaturation 94°C for 5 minutes
- Twenty five cycles of:
  - i. Denaturation at 94°C for 40 seconds

ii. Annealing at 58°C for 40 seconds
iii. Extension at 72°C for 1 minute

Electrophoresis: Mini 6% polyacrylamide gels at 150V for 40 minutes
Staining: Silver nitrate
Result: Fig. 14.3.

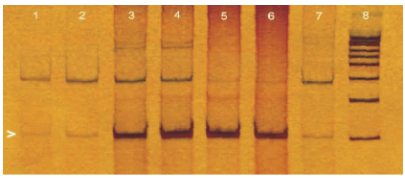


Fig. 14.3. Silver stained PAGE of PCR for JAK-2 mutation. Arrow pointing at the lanes 3-6 show positive while the lanes 1,2 & 7 show negative result.

Bibliography

1. Potter MN, Steward CG, Oakhill A (1993) The significance of detection of minimal residual disease in childhood acute lymphoblastic leukaemia. Br J Haematol 83: 412-418.
2. Khalil SH, Siegrist K, Akhtar M (1997) Application of Polymerase Chain Reaction to detect rearrangement of Immunoglobulin heavy chain genes in lymphoproliferative disease. Ann Saudi Med 17: 395-398.
3. GleißRieder H, Thiel E, Fonatsch C, Janssen LAJ, Heinze B, Janssen JWG, Schoch C, Goekbuget N, Maurer J, Hoelzer D, Bartram CR (2001) Prospective BCR-ABL analysis by polymerase chain reaction (RT-PCR) in adult acute B-lineage lymphoblastic leukemia: reliability of RT-nested-PCR and comparison to cytogenetic data. Leukaemia 15: 1834-1840.
2. Hostein I, Menard A, Soubeyran I, Eghbali H, Debled M, Gastaldello B, Soubeyran P (2001) A 1-kb Bcl-2-PCR fragment detection in a patient with follicular lymphoma and development of a new diagnostic PCR method. Diagn Mol Pathol 10: 89-94.
3. Gabert J, Beillard E, van der Velden VHJ, Bi W, Grimwade D, Pallisgaard N, Barbany G, Cazzaniga G, Cayuela JM, Cave H, Pane F, Aerts JLE, De Micheli D, Thirion X, Pradel V, Gonza´lez M, Viehmann S, Malec M, Saglio G, van Dongen JJM (2003) Standardization and quality control studies of 'real-time' quantitative reverse transcriptase polymerase chain reaction of fusion gene transcripts for residual disease detection in leukemia –A Europe Against Cancer Program. Leukemia 17: 2318–2357.
4. Baxter EJ, Scott LM, Campbell PJ, et al. (2005) Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. Lancet 365: 1054-1061.
5. Krypuy M, Ahmed AA, Etemadmoghadam D, Hyland SJ, Australian Ovarian Cancer Study Group, DeFazio A, Fox SB, Brenton JD, Bowtell DD, Dobrovic A (2007) High resolution melting for mutation scanning of TP53 exons 5–8. BMC Cancer 7: 168-181.
6. Takano EA, Mitchell G, Fox SB, Dobrovic A (2008) Rapid detection of carriers with BRCA1 and BRCA2 mutations using high resolution melting analysis. BMC Cancer 8: 59-66.

15. PCR Applications in Infectious Diseases

PCR based method for detection of pathogens is very sensitive. In a clinical sample targets as small as few molecules of DNA or RNA can be detected. PCR is particularly useful in diagnosis of tuberculosis where culture takes long time or leprosy where culture may not be possible. It is also used for rapid diagnosis of life threatening common bacterial infections and for quick identification of the agents of bioterrorism like Bacillus anthracis and Yersinia pestis. Bacterial antibiotic resistance genes like MRSA (Staph aureus) and VRE (enterococcus), and MDR tuberculosis (katG and rpoB genes) can also be identified by PCR.

PCR is now the method of choice for identification of viral genomes like hepatitis B & C, EBV and HIV etc. An interesting application in viral diseases is in-situ PCR. The virus particles, for example hepatitis-B virus in the liver cells, CMV in lung, and EBV in association with lymphoma, and HPV in cervical cancer can be demonstrated in a smear or tissue specimen. PCR is also being used for many fungal and parasitic infections.

Points to remember in using PCR for infectious agents:

1. With the passage of time the genomes of infectious agents tend to develop subtle changes called polymorphisms. The development of DNA or RNA polymorphism may make a previously designed primer or probe ineffective. Similarly all sub-species/types of an organism or virus may not be detected by the primer or probe due to DNA polymorphism within the species. However, most infectious agents have some conserved regions of the genome. While designing a primer/probe it is important to choose only the conserved regions of the genome.

2. The presence of cross-homology (similarity) between DNA or RNA of the infectious agent and the host genome can give false positive result. More commonly cross-homology may be present between genome of the infectious agent and the organisms that may be normally present in the clinical sample. The problem of cross-homology is best addressed by BLAST searching of the primer/probe before putting them in use (Chapter 1 & 3).

3. PCR would give a positive result whether the organism in the clinical sample is dead or alive. For example a patient on anti-tuberculosis treatment may still be having dead mycobacteria in a clinical sample that could give a positive result on PCR.

4. PCR in itself is a very sensitive technique. But the threshold of detection of a bacterial or viral genome in a clinical sample is largely dependent on the method of extraction of DNA or RNA from the sample. The inconsistency in nucleic acid extraction from the infectious agents may be overcome by using good quality commercial kits made for this purpose.

5. Transport and storage of clinical samples for PCR is also critical. The samples with organisms/viruses where PCR would be done on DNA are

fairly stable under ordinary transport and storage conditions used for culture of such agents. However, clinical samples containing RNA viruses are highly prone to degradation by RNases in the environment. Such samples must be processed as early as is possible. When delay is unavoidable the samples like serum/plasma must be separated and stored at -20°C or below. Repeated freeze thawing of the samples is also damaging for the RNA.

6. If the organism in a clinical sample is low in number it may initially be grown for a short time in a culture medium before extracting its DNA.

There are numerous applications of PCR in the diagnosis of infectious disorders. But here only selected protocols involving common pathogens are described as examples.

PCR for Mycobacterium tuberculosis

DNA extraction

In order to overcome inconsistencies in extraction of microbial DNA it is advisable to use good quality commercial kits.

Real time PCR

A real time PCR for Mycobacterium tuberculosis is described. The PCR primers and the TaqMan® probe (Fig. 15.1) for the insertion sequence element IS986/IS6110 of Mycobacterium tuberculosis CDC1551 (GenBank accession: AE000516.2) were designed by the Primer Express® software (Applied Biosystems, USA).

```
CGGAGCTGCGCGATGGCGAACTCAAGGAGCACATCAGCCGCGTCCACGCCGCCAAC
TACGGTGTTTACGGTGCCCGCAAAGTGTGGCTAACCCTGAACCGTGAGGCATCGAG
GTGGCCAGATGCACCGTCGAACGGCTGATGACCAAACTCGGCCTGTCCGGGACCAC
CCGCGGCAAAGCCCGCAGGACCACGATCGCTGATCCGGCCACAGCCCGTCCCGCCG
ATCTCGTCCAGCGCCGCTTCGGACCACCAGCACCTAACCGGCTGTGGGTAGCAGAC
```

Fig. 15.1. The primers and the TaqMan® probe for the Mycobacterium tuberculosis insertion sequence (IS6110): (GenBank accession: AE000516.2).

Primers:

Forward primer:	5'-GAACCGTGAGGGCATCGA
Reverse primer:	5'-ACAGGCCGAGTTTGGTCATC
TaqMan® Probe:	6 FAM 5'-CCAGATGCACCGTCGAACGGC-BHQ1
Amplified product:	64bp
Primer mix:	Make a mix of the two primers and the probe at concentration of 5pmol each (Chapter 3).

PCR protocol (per sample):

Reaction volume:	25µl
PCR mix:	20µl
Primer mix:	1µl
Taq polymerase:	0.5 units (0.1µl)
DNA:	4µl

Thermal cycling

- Initial denaturation 95°C for 10 minutes
- Forty cycles each of:
  - Denaturation at 95°C for 15 seconds
  - Annealing/extension at 60°C for 1 minute
- Take fluorescence reading at 60°C step

Endpoint PCR for Mycobacterium tuberculosis

The forward and the reverse primers without the probe can also be used for conventional endpoint PCR of Mycobacterium tuberculosis. The amplification is done as described above except that the probe is omitted from the primer mix. At the end of the PCR the amplified product may be run on 2% agarose at 100 volts for 30 minutes or 6% mini-polyacrylamide gel at 150 volts for 15 minutes (Chapter 5). Positive result is shown by 64bp amplified fragment.

PCR for Hepatitis B virus

DNA extraction

For consistent results the viral DNA is best extracted by silica based micro column commercial extraction kits.

Real time PCR

A TaqMan® probe based real time PCR for HBV targeting the conserved 5'Untranslated region of the virus (Fig. 15.2) is described.

TCAATCTTCTCGAGGACTGGGGACCCTGCACCGAACATGGAGAGCACAACATCAGG
ATTCCTAGGACCCCTGCTCGTGTTACAGGCGGGGTTTTTCTTGTTGACAAGAATCC
TCACAATACCACAGAGTCTAGACTCGTGGTGGACTTCTCTCAATTTTCTAGGGGGA
GCACCCACGTGTCCTGGCCAAAATTCGCAGTCCCCAACCTCCAATCACTCACCAAC

Fig. 15.2. Real time PCR primers and TaqMan® probe for the conserved 5'-Untranslated region of Hepatitis-B virus DNA (GenBank accession: NC_003977.1).

Primers and probe:

Forward primer:	5'- GGACCCCTGCTCGTGTTACA
Reverse primer:	5'- GAGAGAAGTCCACCMCGAGTCTAGA
Probe:	6 FAM 5'- TGTTGACAARAATCCTCACCATACCRCAGA-BHQ1 (Reporter: 6 FAM, Quencher: BHQ1)
Primer mix:	Make a mix of the two primers and the probe at concentration of 5pmol each (Chapter 3).

PCR protocol (per sample):

Reaction volume:	25µl
PCR mix:	20µl
Primer mix:	1µl
Taq polymerase:	0.5 units (0.1µl)
DNA:	4µl

Thermal cycling

- Initial denaturation 95°C for 10 minutes
- Forty cycles each of:
  - Denaturation at 95°C for 15 seconds
  - Annealing/extension at 60°C for 1 minute
- Take fluorescence reading at 60°C step

Endpoint PCR for HBV:

The forward and the reverse primers without the probe can also be used for conventional endpoint PCR of HBV. The amplification is done as described above except that the probe is omitted from the primer mix. At the end of the PCR the amplified product may be run on 2% agarose at 100 volts for 30 minutes or 6% mini-polyacrylamide gel at 150 volts for 15 minutes (Chapter 5). Positive result is shown by 90bp amplified fragment.

PCR for Epstein-Barr virus (EBV)

DNA extraction

The viral DNA is best extracted by silica based micro column commercial extraction kits.

Real time PCR

A TaqMan® probe based real time PCR for EBV targeting a conserved region of the virus is described (Fig. 15.3).

GCCCCGTCACGGTGACGTAGTCTGTCTTGAGGAGATGTAGACTTGTAGACACTGC
A**AAACCTCAGGACCTACGCTGC**CC**TAGAGGTTTTGCTAGGGAGGAGACGTGTG**TG
GCTGTAGCCACCCGTCCCGGGTACAAGTCCCGG**GTGGTGAGGACGGTGTCTGT**GG
TTGTCTTCCCAGACTCTGCTTTCTGCCGTCTTCGGTCAAGTACCAGCTGGTGGTC

Fig. 15.3. DNA sequence of Epstein-Barr virus (EBV) genome (GenBank accession: V01555.2) showing the real time PCR primers and the TaqMan® probe.

Primers:

Forward primer:	5'-AAACCTCAGGACCTACGCTGC
Reverse primer:	5'-ACAGACACCGTCCTCACCAC
Probe:	FAM 5'-TAGAGGTTTTGCTAGGGAGGAGACGTGTG-BHQ1 (Reporter: 6 FAM, Quencher: BHQ1)
Primer mix:	Make a mix of the two primers and the probe at concentration of 5pmol each (Chapter 3).

PCR protocol (per sample):

Reaction volume:	25µl
PCR mix:	20µl
Primer mix:	1µl
Taq polymerase:	0.5 units (0.1µl)
DNA:	4µl

Thermal cycling

- Initial denaturation 95°C for 10 minutes
- Forty cycles each of:
- Denaturation at 95°C for 15 seconds
- Annealing/extension at 60°C for 1 minute
- Take fluorescence reading at 60°C step

Endpoint PCR

The forward and the reverse primers without the probe can also be used for conventional endpoint PCR of EBV. The amplification is done as described above except that the probe is omitted from the primer mix. At the end of the PCR the amplified product may be run on 2% agarose at 100 volts for 30 minutes or 6% mini-polyacrylamide gel at 150 volts for 15 minutes (Chapter 5). Positive result is shown by 107bp amplified fragment.

PCR for Cytomegalo Virus (CMV)

DNA extraction

For consistent results the viral DNA is best extracted by silica based micro column commercial extraction kits.

Real time PCR

A TaqMan® probe based real time PCR for CMV targeting the human cytomegalovirus glycoprotein B gene of the virus is described (Fig. 15.4).

```
CTTCTTCCCAAACGGTCAGCCATGGTGTTAACGAGACCATCTACAACACTACCCT
CAAGTACGGAGATGTGGTGGGGGTCAATACCACCAAGTACCCCTATCGCGTGTGT
TCTATTGGCCCAGGGTACGGATCTTATTCGCTTTGAACGTAATATCGTCTGCACCT
CGATGAAGCCCATCAATGAAGACCTGGACGAGGGCATCATGGTGGTCTACAAACG
CAACATCGTCGCGCACACCTTTAAGGTACGAGTCTACCAGAAGGTTTTGACGTTT
```

Fig. 15.4. DNA sequence of Cytomegalovirus glycoprotein B gene (GenBank accession: M60929.1) showing the real time PCR primers and the TaqMan® probe.

Primers:

Forward primer:	5'-AAGTACCCCTATCGCGTGTG
Reverse primer:	5'-ATGATGCCCTCGTCCAGGTC
Probe:	6 FAM 5'-TGGCCCAGGGTACGGATCTTATTCG-BHQ1 (Reporter: 6 FAM, Quencher: BHQ1)
Primer mix:	Make a mix of the two primers and the probe at concentration of 5pmol each (Chapter 3).

PCR protocol (per sample):

Reaction volume:	25µl
PCR mix:	20µl
Primer mix:	1µl
Taq polymerase:	0.5 units (0.1µl)
DNA:	4µl

Thermal cycling

- Initial denaturation 95°C for 10 minutes
- Forty cycles each of:
  - Denaturation at 95°C for 15 seconds
  - Annealing/extension at 60°C for 1 minute
- Take fluorescence reading at 60°C step

Endpoint PCR

The forward and the reverse primers without the probe can also be used for conventional endpoint PCR of CMV. The amplification is done as described above except that the probe is omitted from the primer mix. At the end of the PCR the amplified product may be run on 2% agarose at 100 volts for 30 minutes or 6% mini-polyacrylamide gel at 150 volts for 15 minutes (Chapter 5). Positive result is shown by 116bp amplified fragment.

PCR for HCV

PCR of HCV, being RNA, is done in two steps. In the first step HCV RNA is converted to cDNA and in the second step cDNA is amplified by PCR. A method targeting the highly conserved 5'-untranslated region of the virus is described. Fig. 15.5 shows a set of primers and TaqMan® probes for the 5' untranslated region of HCV genome.

CACTCCACCATGAATCACTCCCCTGTGAGGAACTACTGTCTTCACGCAGAA**AGCGT**
CTAGCCATGGCGTTAGTATGAGTGTCGTGCAGCCTCCAGGAC**CCCCCCTCCCGGGA**
GAGCCATAGTGGTCTGCGGAACCGGT**GAGTACACCGGAATTGCCAGGA**CGACCGGG
TCCTTTCTTGGATCAACCCGCTCAATGCCTGGAGATTTGGGCGTGCCCCCGCAAGA

Fig. 15.5. Sequence of 5' untranslated region of HCV genome (GenBank: EU678744.1) showing a set of primers and TaqMan® pobe for real time PCR.

RNA Extraction

For best and consistent results RNA extraction by commercial silica based extraction columns is recommended.

Preperation of cDNA:

Reverse transcription (RT):

Primer (HCV-R) 5'-TCCTCGCAATTCCGGTGTACTC

Reactions conditions:

- 5 X reaction buffer: 4 µl
- 10 mmol dNTPs mix: 2 µl
- RT Primer (10 pmol/µl): 1 µl
- MMLV- Reverse Transcriptase (200 U/µl): 1 µl
- RNase inhibitor (20 U/µl): 1 µl
- Deionized water: 11 µl
- RNA: 2 µl
- Incubation: 42°C for 60 minutes
- RT inactivation: 70°C for 5 minutes

•

•Real Time PCR

PCR Primers:

HCV-F	5'-AGCGTCTAGCCATGGCGTTAGTAT
HCV-R	5'- TCCTCGCAATTCCGGTGTACTC
HCV Probe	6 FAM-CCCCCCTCCCGGGAGAGCCATAGT-BHQ1
Amplified Product:	109bp
Reaction volume:	25µl
PCR mix:	21µl
Primer mix:	1µl (containing each primer at 5pmol/µl concentration)
Taq polymerase:	0.5 units (0.1µl)
cDNA:	3µl

Thermal cycling

- Initial denaturation 95°C for 10 minutes
- Forty cycles each of:
  - Denaturation at 95°C for 15 seconds
  - Annealing/extension at 60°C for 1 minute
- Take fluorescence reading at 60°C step

Endpoint PCR for HCV

The forward and the reverse primers without the probe can also be used for conventional endpoint PCR of HCV. cDNA is prepared as described above using the HCV-R primer. The amplification is also done as described above except that the probe is omitted from the primer mix. At the end of the PCR the amplified product may be run on 2% agarose at 100 volts for 30 minutes or 6% mini-polyacrylamide gel at 150 volts for 15 minutes (Chapter 5). Positive result is shown by a 109bp amplified fragment.

PCR for Dengue virus:

A pan dengue PCR is described that can identify all four genotypes of the virus. PCR of Dengue virus, being RNA, is done in two steps. In the first stem viral RNA is converted to cDNA and in the second step cDNA is amplified by PCR. A method using the primers and TaqMan® MGB probe (Fig. 15.6) for the highly conserved 3' untranslated region shared by the four genotypes of the Dengue virus genome is described. The TaqMan® MGB probe is conjugated with a minor groove binder (MGB) at the 3' end that allows higher probe Tm with a relatively short length of the probe.

DENV-1 (GenBank NC_001477.1)

AGACACAACGCAGCAGCGGGGCCCAACACCAGGGGAAGCTGTACCCTGGTGGTAAGGACTA
GAGGTTAGAGGAGACCCCCCGCACAACAACAAACAGCATATTGACGCTGGGAGAGACCAGA
GATCCTGCTGTCTCTACAGCATCATTCCAGGCACAGAACGCCAAAAAATGGAATGGTGCTG
TTGAATCAACAGGTTCT

DENV-2 (GenBank *AF038403.1*)

CAAATCGCAGCAACAATGGGGGCCCAAGGTGAGATGAAGCTGTAGTCTCACTGGAAGGACT
AGAGGTTAGAGGAGACCCCCCCAAAACAAAAACAGCATATTGACGCTGGGAAAGACCAGA
GATCCTGCTGTCTCCTCAGCATCATTCCAGGCACAGAACGCCAGAAAATGGAATGGTGCTG
TTGAATCAACAGGTTCT

DENV-3 (GenBank *NC_001475.2*)

ATGACACAACGCAGCAGCGGGGCCCGAGCACTGAGGGAAGCTGTACCTCCTTGCAAAGGAC
TAGAGGTTAGAGGAGACCCCCCGCAAATAAAAACAGCATATTGACGCTGGGAGAGACCAGA
GATCCTGCTGTCTCCTCAGCATCATTCCAGGCACAGAACGCCAGAAAATGGAATGGTGCTG
TTGAATCAACAGGTTCT

DENV-4 (GenBank *NC_002640.1*)

GATAAAACGCAGCAAAAGGGGGCCCGAAGCCAGGAGGAAGCTGTACTCCTGGTGGAAGGAC
TAGAGGTTAGAGGAGACCCCCCCAACACAAAAACAGCATATTGACGCTGGGAAAGACCAGA
GATCCTGCTGTCTCTGCAACATCAATCCAGGCACAGAGCGCCGCAAGATGGATTGGTGTTG
TTGATCCAACAGGTTCT

Fig. 15.6. The primers and the TaqMan® MGB probe for real time PCR targeting the 3' un-translated region of four genotypes of Dengue virus. Minor differences in the sequence of the four dengue virus genotypes are shown as green.

RNA Extraction

For best and consistent results RNA extraction by commercial silica based extraction columns is recommended.

Preperation of cDNA:

Reverse transcription (RT):
Primer (DENV-R) 5'-ACCATTCCATTTTCTGGCGTT
Reactions conditions:

- 5 X reaction buffer: 4 µl
- 10 mmol dNTPs mix: 2 µl
- RT Primer (10 pmol/µl): 1 µl
- MMLV- Reverse Transcriptase (200 U/µl): 1 µl
- RNase inhibitor (20 U/µl): 1 µl
- Deionized water: 11 µl
- RNA: 2 µl
- Incubation: 42°C for 60 minutes
- RT inactivation: 70°C for 5 minutes

Real Time PCR

PCR Primers:

DENV-F	5'- GARAGACCAGAGATCCTGCTGTCT
DENV-R	5'- ACCATTCCATTTTCTGGCGTT
DENV-Pro	FAM 5'-AGCATCATTCCAGGCAC-BHQ1-MGB
Amplified Product:	68bp
Reaction volume:	25µl
PCR mix:	21µl
Primer mix:	1µl (containing each primer at 5pmol/µl concentration)
Taq polymerase:	0.5 units (0.1µl)
cDNA:	3µl

Thermal cycling

- Initial denaturation 95°C for 10 minutes
- Forty cycles each of:
  - Denaturation at 95°C for 15 seconds
  - Annealing/extension at 60°C for 1 minute
- Take fluorescence reading at 60°C step

Endpoint PCR for Dengue virus

The forward and the reverse primers without the probe can also be used for conventional endpoint PCR of Dengue virus. cDNA is prepared as described above using the DENV-R primer. The amplification is also done as described above except that the probe is omitted from the primer mix. At the end of the PCR the

amplified product may be run on 2% agarose at 100 volts for 30 minutes or 6% mini-polyacrylamide gel at 150 volts for 15 minutes (Chapter 5). Positive result is shown by a 68bp amplified fragment.

PCR for Malaria

A TaqMan® probe based real time PCR for malaria targeting a conserved region of Plasmodium falciparum, P. vivax, P. malariae, and P. ovale is described (Fig. 15.7).

P. falciparum (GenBank M19172.1)

TAGTGTGTATCAATCGAGTTTCTGACCTATCAGCTTTTGATGTTAGGGTATTGGCCT
A**ACATGGCTATGACGGGTAACG**GGGAATTAGAGT**TCGATTCCGGAGAGGGAGCCTGA**
GAAA**TAGCTACCACATCTAAGGAAGGCA**GCAGGCGCGTAAATTACCCAATTCTAAAG
AAGAGAGGTAGTGACAAGAAATAACAATGCAAGGCCAATTTTTGGTTTTGTAA

P. vivax (GenBank X13926.1)

TTAGTGTGTATCAATCGAGTTCTGACCTATCAGCTTTTGATGTTAGGGTATTGGCCT
A**ACATGGCTATGACGGGTAACG**GGGAATTAGAGT**TCGATTCCGGAGAGGGAGCCTGA**
GAAA**TAGCTACCACATCTAAGGAAGGCA**GCAGGCGCGTAAATTACCCAATTCTAAAG
AAGAGAGGTAGTGACAAGAAATAACAATACAAGGCCAATCTGGCTTTGTAATT

P. malariae (GenBank M54897.1)

TAGTGTGTATCAATCGAGTTTCTGACCTATCAGCTTTTGATGTTAGGGTATTGGCCT
A**ACATGGCTATGACGGGTAACG**GGGAATTAGAGT**TCGATTCCGGAGAGGGAGCCTGA**
GAAA**TAGCTACCACATCTAAGGAAGGCA**GCAGGCGCGTAAATTACCCAATTCTAAAG
AAGAGAGGTAGTGACAAGAAATAACAATGCAAGGCCAAATTTTGGTTTTGCAA

P. ovale (GenBank: L48987.1)

TAGTGTGTATCAATCGAGTTTCTGACCTATCAGCTTTTGATGTTAGGGTATTGGCCT
A**ACATGGCTATGACGGGTAACG**GGGAATTAGAGT**TCGATTCCGGAGAGGGAGCCTGA**
GAAA**TAGCTACCACATCTAAGGAAGGCA**GCAGGCGCGTAAATTACCCAATTCTAAAG
AAGAGAGGTAGTGACAAGAAATAACAATACAAGGCCATTTCATGGTTTTGTAA

Fig. 15.7. Sequence and the primers for real time PCR targeting the Small Subunit (SSU) RNA gene of the four subspecies of malarial parasite.

DNA extraction:

Commercial or Chelex based method on whole blood collected in EDTA (Chapter 2). Chelex based method may give increased background fluorescence.

Real Time PCR

Forward primer:	5'-ACATGGCTATGACGGGTAACG
Reverse primer:	5'-TGCCTTCCTTAGATGTGGTAGCTA
Probe:	6 FAM 5'-TCAGGCTCCCTCTCCGGAATCGA-BHQ1 (Reporter: 6 FAM, Quencher: BHQ1)
Primer mix:	Make a mix of the two primers and the probe at concentration of 5pmol each (Chapter 3).

PCR protocol (per sample):

Reaction volume:	25µl
PCR mix:	22µl
Primer mix:	1µl
Taq polymerase:	0.5 units (0.1µl)
DNA:	2µl

Thermal cycling

t. Initial denaturation 95°C for 10 minutes
u. Forty cycles each of:
    i. Denaturation at 95°C for 15 seconds
    ii. Annealing/extension at 60°C for 1 minute
v. Take fluorescence reading at 60°C step

Endpoint PCR

The forward and the reverse primers without the probe can also be used for conventional endpoint PCR of Malaria. The amplification is done as described above except that the probe is omitted from the primer mix. At the end of the PCR the amplified product may be run on 2% agarose at 100 volts for 30 minutes or 6% mini-polyacrylamide gel at 150 volts for 15 minutes (Chapter 5). Positive result is shown by 84bp amplified fragment.

PCR for Leishmania donovani

A TaqMan® probe based real time PCR for Leishmania donovani glucosephosphate isomerase gene (Fig.15.8) is described.

GTACGGCAGCACCGCCTCTGTCTCCGCGCCGAAAAAGTTGTTGTA **CCAGATGCCGA**
CCAAAGCCAGCATC**ATCGGCAGGTTCTGCTCCGTCG**GTGCAGACGCAAAGTG**GTTA**
TCCATCACGTGCGCGCCAGTCAGGAACTCCACAAAGTTGTCGTAGCCGATCGAAAG
CATCACGGAGAGACCGATGGCGGACCACACAGAGTAGCGACCACCGACCCAGTCCC

Fig. 15.8. DNA sequence of Leishmania donovani isolate WR 378 glucosephosphate isomerase gene (GenBank: AY974201.1).

DNA extraction

DNA extraction by phenol chloroform or chelex method from 0.2ml bone marrow aspirate or archival bone marrow smears (Chapter 2).

Real Time PCR

Forward primer:	5'-CCAGATGCCGACCAAAGC
Reverse primer:	5'-CGCGCACGTGATGGATAAC
Probe:	6 FAM 5'-ATCGGCAGGTTCTGCTCCGTCG-BHQ1 (Reporter: 6 FAM, Quencher: BHQ1)
Primer mix:	Make a mix of the two primers and the probe at concentration of 5pmol each (Chapter 3).

PCR protocol (per sample):

Reaction volume:	25µl
PCR mix:	22µl
Primer mix:	1µl
Taq polymerase:	0.5 units (0.1µl)
DNA:	2µl

Thermal cycling

- Initial denaturation 95°C for 10 minutes
- Forty cycles each of:
  - Denaturation at 95°C for 15 seconds
  - Annealing/extension at 60°C for 1 minute
- Take fluorescence reading at 60°C step

Endpoint PCR for Leishmania donovani

The forward and the reverse primers without the probe can also be used for conventional endpoint PCR of Leishmania donovani. The amplification is done as described above except that the probe is omitted from the primer mix. At the end of the PCR the amplified product may be run on 2% agarose at 100 volts for 30 minutes or 6% mini-polyacrylamide gel at 150 volts for 15 minutes (Chapter 5). Positive result is shown by a 82bp amplified fragment.

Bibliography

1. Lee MA, Tan CH, Aw LT, Tang CS, Singh M, Lee SH, Chia HP, Yap EPH (2002) Real-Time Fluorescence-Based PCR for Detection of Malaria Parasites. J Clin Microbiol 40: 4343–4345.
2. Jebbink J, Bai X, Rogers BB, Dawson DB, Scheuermann RH, Domiati-Saad R (2003) Development of Real-Time PCR Assays for the Quantitative Detection of Epstein-Barr Virus and Cytomegalovirus, Comparison of TaqMan Probes, and Molecular Beacons. J Mol Diagn 5: 15–20.

3. Wortmann G, Hochberg L, Houng HH, Sweeney C, Zapor M, Aronson N, Weina P, Ockenhouse CF (2005) Rapid identification of Leishmania complexes by real time PCR assay. Am J Trop Med Hyg 73: 999–1004.
4. Espy MJ, Uhl JR, Sloan LM, Buckwalter SP, Jones MF, Vetter EA, Yao JDC, Wengenack NL, Rosenblatt JE, Cockerill III FR, Smith TF (2006) Real-Time PCR in Clinical Microbiology: Applications for routine laboratory testing. Clin Microbiol Rev 19: 165–256.
5. Pas SD, Fries E, DeMan RA, Osterhaus ADME, Niesters HGM (2000) Development of a quantitative real-time detection assay for hepatitis B virus DNA and comparison with two commercial assays. J Clin Microbiol 38: 2897-2901.
6. Meng S and Li J (2010) A novel duplex real-time reverse transcriptase-polymerase chain reaction assay for the detection of hepatitis C viral RNA with armored RNA as internal control. Virology Journal 7: 117-125.
7. Gurukumar KR, Priyadarshini D, Patil JA, Bhagat A, Singh A, Shah PS, Cecilia D (2009) Development of real time PCR for detection and quantitation of Dengue viruses. Virology Journal 6: 10-17.
8. Wortmann G, Hochberg L, Houng H, Sweeney C, Zapor M, Aronson N, Weina P, Ockenhouse CF (2005) Rapid identification of Leishmania complexes by a real-time PCR assay. Am J Trop Med Hyg 73: 999-1004.

16. Forensic DNA Testing

DNA from no two individuals except identical twins is alike. The person to person differences in DNA can be discovered by PCR amplification and genomic sequencing. The main advantages of DNA typing are the universality of application, an almost unlimited power to discriminate, extreme sensitivity, and reasonable resistance to degradation by environmental factors. The typing of DNA can be used in:

- Linking a suspect to a crime
- Excluding a falsely accused person
- Recognizing serial crimes
- Resolving parentage disputes
- Identification of the remains of victims

The differences in DNA are either in the form of variable number of tandem repeats (VNTR) or single nucleotide polymorphisms (SNP). Approximately 20% of the human genome comprises tandem repeat sequences. These are called micro-satellites or short tandem repeats (STR) when the repeats are 2-6bp in length. When the repeat sequences are 7-80bp in length these are called mini-satellite or variable number of tandem repeats (VNTR). The number of STR repeat units tends to vary between individuals. This variation (polymorphism) makes them extremely useful in applications like human identification and linkage analysis for diagnosis of genetic disorders.

STR analysis:

There are over a million sites (loci) in the human genome that have STRs of which over 20,000 have been characterized. Considering the highly polymorphic nature and the ease with witch the STRs can be analyzed these provide an extremely powerful tool for human identification. The discrimination power of STRs increases with increase in the number of loci tested. Therefore it is usual to use a battery of STRs. The repeat sequence of STRs range from 2-6bp. Due to technical reasons four base pair (tetra-nucleotide) repeats are best suited for forensic case work.

STR loci

The STRs are mostly located in the non-coding DNA between the genes (inter-genic DNA). Some STRs are also present in the intervening sequences (introns) of known genes. The STRs found in the inter-genic DNA are named according to the chromosome number and the site. For example in the name “D5S818” “D” stands for DNA, “5” is the chromosome number, “S” stands for single copy sequence and “818” is the locus number. The STRs in the introns of the known genes are identified by their location e.g. TH01 is present in the intron-1 of human tyrosine hydroxylase gene and TPOX is located in the thyroid peroxidase gene.

In a typical tetra-nucleotide STR the repeat units comprise four base pairs. For example an STR with “GATA” repeat would look like:

GTATCCTTATGTAATATTTTGAAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGGTAGATAGAGGTATAAATAAGGATACAGATATAG

Autosomal STRs

In most forensic DNA applications STRs located on autosomes are used. A list of 15 core STR loci and their chromosomal locations are shown in Table 16.1 (Short Tandem Repeat DNA Internet DataBase).

Table. 16.1. Core STR loci used in forensic DNA testing.

Name	Chromosome	Location	Type	GenBank
TPOX	2p23-2pter	Intron 10 of human thyroid peroxidase gene	Simple	M68651
D3S1358	3p21	-	Complex	11449919
FGA	4q28	Intron 3 of alpha fibrinogen gene	Complex	M64982
D5S818	5q21-q31	-	Simple	G08446 AC008512
CSF1PO	5q33.3-34	c-fms proto-oncogene for CSF-1 receptor gene	Simple	X14720
D7S820	7q	-	Simple	G08616 AC004848
D8S1179	8q24.1-24.2	-	Simple	G08710 AF216671
Th01	11p15-15.5	Intron 1 of human tyrosine hydroxylase gene	Complex	D00269
vWA	12p12-pter	Gene for von Willebrand antigen.	Complex	M25858
D13S317	13q22-q31	-	Simple	G09017 AL353628.2
D16S539	16q22-24	-	Simple	G07925 AC024591.3
D18S51	18q21.3	-	Simple	X91254 AP001534
D21S11	21q21.1	-	Complex	M84567 AP000433
D2S1338	2q35	-	Complex	AC010136 G08202
D19S433	19q12	-	Complex	G08036 AC008507.6

Y-Chromosome STRs

Several STRs have also been identified on Y-chromosome. These are useful in tracing male DNA in investigations like sexual assault. These may also be used in investigation of paternal lineage inheritance. A list of 15 Y-STR markers available as a commercial kit (Applied Biosystems) is shown in Table 16.2 (Short Tandem Repeat DNA Internet DataBase).

Table. 16.2. Fifteen Y-STRs available as a commercial kit (Applied Biosystems, USA)

Locus	Repeat Numbers	Repeat Motif	GenBank Accession	Reference Allele
DYS 19	10-19	TAGA	AC017019	15
DYS 385 a/b	7-28	GAAA	AC022486	11
DYS 389 I DYS 389 II	DYS 389I: 9-17 DYS 389II:24-34	(TCTG) (TCTA) (TCTG) (TCTA)	AC004617	12, 29
DYS 390	17-28	(TCTA) (TCTG)	AC011289	24
DYS 391	6-14	TCTA	AC011302	11
DYS 392	6-17	TAT	AC011745	13
DYS 393	9-17	AGAT	AC006152	12
DYS 437	13-17	TCTA	AC002992	16
DYS 438	6-14	TTTTC	AC002531	10
DYS 439	9-14	AGAT	AC002992	13
DYS 448	20-26	AGAGAT	AC025227	22
DYS 456	13-18	AGAT	AC010106	15
DYS 458	13-20	GAAA	AC010902	16
DYS 635	17-27	TSTA compound	AC004772	23

STR allele nomenclature

The STR alleles are named according to the number of repeat units it contains e.g. 7, 8, 9, 10, 11 etc. Some STRs are more complex than simple repeats. The complexity may be present both in the sequence as well as the number of bases in the repeat unit. The variation in sequence of the repeats can be found only by genomic sequencing. However, the variation in number of bases in a repeat results in different sizes that can be picked on gel electrophoresis. For example at the TH01 locus allele 9 has nine simple repeats i.e. [AATG]9. Another allele at the same locus has an additional triplet i.e. [AATG]6ATG[AATG]3. The resulting allele is 3bp longer and is written as 9.3. The D21S11 locus contains numerous complex alleles like 32.1, 32.2 etc. The complex loci being more polymorphic are more informative.

The STRs are inherited in a simple Mendelian fashion. An individual inherits an allele each from its father and the mother. A person may be homozygous (the same allele on the maternal and the paternal chromosomes) or compound heterozygous (different alleles on the two chromosomes). A typical STR profile comprising genotypes at several loci is shown in Table 16.3.

Table 16.3. The STR profile at various loci in individuals A, B, and C.

Individuals	D3S1358	D5S818	D7S820	D8S117	D21S11	TH01	TPOX
A.	15,18	12,12	9,10	12,16	28,29	7,9	9,11
B.	14,14	10,12	8,10	11,15	30,32.2	7,9.3	8,11
C.	15,17	10,10	9,11	13,15	29,32.2	8,10	8,11

Allele frequencies

In a given population the STR allele frequencies are calculated by simple counting. Each individual has two alleles at each locus. Genotyping of 100 individuals from a population would mean examination of 200 chromosomes (alleles). Each allele of a compound heterozygote is counted as one and homozygotes are counted as two. For example if 20/100 people have allele 8, including 2 homozygotes (8,8), the frequency of allele 8 would be 18+2+2 = 22/200 i.e. 11% or 0.11.

Minimum allele threshold

In a population survey the uncommon or the rare alleles are expected to have an under-representation. In order to overcome this problem it has been recommended to inflate the frequency of rare alleles (<5 counts) to 5. The 5/2N formula is used for this purpose where N is the number of individuals examined. The N is doubled because each individual has two chromosomes (alleles). For example if allele 12 is observed in 2/100 people its frequency by the conventional method would be 2/200 = 1% (0.01). However, by the 5/2N formula its frequency would be 2.5% (0.025).

Genotype frequencies

The allele combination in an individual at a locus is called its genotype e.g. "7,10" or "7,7" etc. The number of possible genotypes increases with an increase in the number of alleles. The possible genotypes can be calculated by the formula [n(n+1)/2] where n is the number of different alleles found in the population. The observed genotype frequencies are calculated by simple counting. The expected genotype frequencies can be calculated from the observed allele frequencies by using Hardy Weinberg equation ($p_2 + q_2 + 2pq = 1$). The homozygotes of two alleles with frequencies of p and q would be equal to p_2 and q_2 respectively while the heterozygotes (compound heterozygotes) would be equal to 2pq.

Example

Table 16.4 describes an example of calculation of expected genotype frequencies of alleles 10 and 11 with observed frequencies of p and q respectively.

Table 16.4. The expected genotypes of two alleles at D5S818 locus.

D5S818	Frequency	p^2	q^2	2pq
Allele 10	$p = 0.108$	0.0117	-	0.0687
Allele 11	$q = 0.318$	-	0.101	0.0687

Profile frequencies

The STR profile of an individual is the combination of genotypes at several loci. Larger the number of loci tested rarer would be the combination in the population. The frequency of a profile is calculated by the multiplication rule. The individual genotype frequencies are multiplied to get the combined frequency (see example). The combined frequency of 15 core STR loci in the US Caucasian population is 1 in 2.46 quadrillion (1015).

Example

Table 16.5 gives an example of how a profile frequency can be calculated from various allele and genotype frequencies.

Table 16.5. Example of calculation of profile frequency from the genotype frequencies in a given population.

Loci	Alleles	Allele frequency	p^2	q^2	$2pq$
D3S1358	15	p = 0.299	0.0894	-	0.0466
	18	q = 0.078	-	0.0061	0.0466
D5S818	10	p = 0.108	0.0117	-	0.0687
	11	q = 0.318	-	0.1011	0.0687
D8S1179	12	p = 0.103	0.0107	-	0.0132
	16	q = 0.064	-	0.0041	0.0132
D21S11	30.2	p = 0.059	0.0035	-	-
	30.2	-	-	-	-
Profile		Profile frequency			
D3 & D5		0.0466 X 0.0687 = 0.0032 (1 in 312)			
D3, D5 & D8		0.0466 X 0.0687 X 0.0132 = 0.000042 (1 in 23,809)			
D3, D5, D8 & D21*		0.0466 X 0.0687 X 0.0132 X 0.0035 = 0.00000015 (1 in 6666,666)			

*additional loci can also be added to this calculation

Forensic calculations and consanguineous marriage

Hardy-Weinberg equation for the calculation of expected genotype frequencies is for populations where mating is random. In a population where consanguineous marriage and marriage between tribe members is very common, the population has several substructures (smaller groups). Therefore Hardy-Weinberg equation would not be applicable as such.

The main genetic effect of consanguineous marriage is an increase in the proportion of homozygotes and a corresponding reduction in the hetrozygotes. The increase in homozygotes as compared to Hardy-Weinberg proportions is by an amount Fpq, while heterozygotes are reduced by 2Fpq where F is the inbreeding coefficient, and p and q are the frequencies of the alleles under consideration.

$$\text{Homozygotes} = p_2 + Fpq \text{ or } q_2 + Fpq$$
$$\text{Heterozygotes} = 2pq\,(1 - F)$$

The coefficient of inbreeding is the probability that an individual receives at a given locus two genes that are identical by descent (copies of a single gene carried by a common ancestor). The value of F for a first cousin marriage is 0.0625 i.e. 6.25% of the genes are identical by descent. In 1½ cousin and 2nd cousin marriage F is 0.0313 and 0.0156 respectively. The average coefficient of inbreeding in the Pakistani population is 0.0280 that may be as high as 0.0350 in selected populations/tribes.

Example

The correction for consanguineous marriage applied to the genotype and profile frequency on the example shown in table 16.5 is presented in table 16.6.

Table 16.6. Example of correction for consanguineous marriage applied to the genotype and profile frequency of the example presented in Table 16.5.

Loci	Alleles	Allele frequency	$p^2 + Fpq$	$q^2 + Fpq$	$2pq\ (1 - F)$
D3S1358	15	$p = 0.299$	0.0900	-	0.0453
	18	$q = 0.078$	-	0.0068	0.0453
D5S818	10	$p = 0.108$	0.0127	-	0.0668
	11	$q = 0.318$	-	0.1021	0.0668
D8S1179	12	$p = 0.103$	0.0109	-	0.0128
	16	$q = 0.064$	-	0.0059	0.0128
D21S11	30.2	$p = 0.059$	0.0036*	-	-
	30.2	-	-	-	-
Profile		Profile frequency			
D3 & D5		0.0453 X 0.0668 = 0.00303 (1 in 330)			
D3, D5 & D8		0.0453 X 0.0668 X 0.0128 = 0.000039 (1 in 25,641)			
D3, D5, D8 & D21		0.0453 X 0.0668 X 0.0128 X 0.0036 = 0.0000014 (1 in 7142,857)			

* $p_2 + Fp(1-p)$

Mutations in STRs

Genomic DNA is liable to develop spontaneous mutations with the passage of time. In fact the existence of highly polymorphic STRs in the genome is thought to be due to spontaneous mutations that is a fairly regular event. The STR mutations become significant if these are encountered in solving a case of inheritance.

The rate of spontaneous mutations at the STR loci ranges from 1 in 500 to 1 in 1000. The rates of mutations at the core STR loci used in forensic DNA testing are given in Table 16.7. A mutation may be suspected if a disagreement is found between alleles of the parents and the offspring at one odd locus out of the several tested. The mutation can be confirmed by genomic sequencing. It may be pointed out that matching between DNA samples from a crime scene and a suspect would not be affected by mutations in the STR loci.

Table 16.7 Mutation rates at the STR loci commonly used in forensic DNA testing.

STR Loci	Mutation Rate	STR Loci	Mutation Rate	STR Loci	Mutation Rate
CSF1PO	0.16%	D3S1358	0.12%	D16S539	0.11%
FGA	0.28%	D5S818	0.11%	D18S51	0.22%
TH01	0.01%	D7S820	0.10%	D21S11	0.19%
TPOX	0.01%	D8S1179	0.14%	D2S1338	0.12%
VWA	0.17%	D13S317	0.14%	D19S433	0.11%

Core STR loci used in human identification

The core STR loci are sets of DNA markers that are globally accepted for human identification (Table 16.8). The uniformity is adopted to share and compare genetic information between different labs and the legal systems. The loci have been carefully selected to avoid their linkage with any physical character or genetic disease. In addition to the STRs a marker for sex determination, usually Amelogenin, is also included in the profile.

Table 16.9 Core STR loci required by various countries and the legal systems.

Countries	Core STR Loci
US	CSF1PO, FGA, TH01, TPOX, vWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Amelogenin
UK/European	FGA, TH01, vWA, D2S1338, D3S1358, D8S1179, D16S539, D18S51, D19S433, D21S11, Amelogenin Recommended Loci: D1S1656, D2S441, D10S1248, D12S391, D22S1045, TPOX
Interpol	FGA, TH01, vWA, D3S1358, D8S1179, D18S51, D21S11, Amelogenin

Combined DNA Index System (CODIS)

United States Federal Bureau of Investigation has created a database, called CODIS, that stores the DNA profiles of convicted offenders and the biological material found at crime scenes. The database contains DNA profiles comprising 15 STR loci including CSF1PO, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D19S433, D21S11, FGA, TH01, TPOX, vWA and Amelogenin to determine sex. Inclusion of profiles in the CODIS database is authorized by the DNA Identification Act of 1994. At present over 5 million entries exist in the database making it the largest in the world. It has helped in over 93,000 criminal investigations.

Single Nucleotide Polymorphisms (SNP)

SNPs are DNA sequence variations that are distributed throughout the genome and are found after every 300 to 500 base pairs. Almost all SNPs have two alleles. By convention the base change is called polymorphism when the frequency of its alleles in a population is more than 1%. SNPs are present in the non-coding as well as the coding sequences of DNA.

The SNPs are identified by restriction enzyme analysis and gel electrophoresis (Chapter 8). More recently real time PCR is being used to analyze multiple SNPs in a single tube by multiplex allele specific probes. The automated SNP genotyping by micro arrays also allow analysis of very large number of SNPs on a single gene chip. The later has potential for human identification applications like investigation of mass disasters.

Mitochondrial DNA

Each mitochondrion contains 2-10 copies of a circular piece of DNA, 16,569 base pairs in length. It is inherited from the mother. Mitochondrial DNA contains two hyper variable regions (HV1 and HV2) that contain many SNPs. The regions can be amplified and sequenced. Since mitochondrial DNA is inherited from the mother it can be useful in tracing maternal inheritance. Mitochondrial DNA can be extracted from s like hair shafts, old bones and teeth etc.

PCR for sex determination

Determination of sex is an essential part of human identification by DNA test. Most commonly it is done by amplifying a sequence from the amelogenin gene whose length varies between male and the female.

Degraded DNA and "Mini STRs"

The DNA is a very large molecule that can be easily broken to smaller fragments by shearing force or bacterial enzymes (Chapter 2). Such degraded DNA samples are poorly amplified. The problem is most marked in highly degraded DNA samples. In precious and trace forensic samples degradation can completely jeopardize the analysis.

The problem of degraded DNA can be overcome to some extent by using "Mini STRs". In a mini STR analysis the forward and the reverse amplification primers flanking the repeat units are brought to the nearest possible distance from the repeat units. The resultant amplified fragments are of smallest possible size. Many degraded DNA samples can be analyzed to provide sufficient information.
Limitations of DNA test for human identification
Some of the limitations of forensic DNA test are:

1. It can not tell the age of the person.
2. In some situations it may provide information about predisposition to disease, color of eyes, height or hair color.

Applications

Matching suspect with evidence

The profiles of the two or more DNA samples are aligned to see if there is any difference or not. There could be three possible outcomes:

1. Match: When the two or more samples have the same genotypes. The statistical significance of the match is calculated as described in a subsequent section.
2. Exclusion: When the two samples originate from two different sources. This does not require prior knowledge of the allele frequencies in the population.
3. Inconclusive: Neither of the above two outcomes.

Example

The STR profiles of four DNA samples are shown in Table 16.10. Sample 1 was picked from crime scene whereas samples 2-4 are of the suspects. The samples 1 and 2 are completely matching whereas the samples 3 and 4 do not match and are therefore excluded.

Table 16.10 Comparison of the STR profile of a DNA sample collected from a crime scene (serial 1) and three suspects (serial 2-4). The profile of sample at serial 2 completely matches with that of the crime scene DNA whereas the samples at serial 2 & 3 are excluded from the match.

Sample	D2S1338	D3S1358	D5S818	D7S820	D8S1179	D13S317	D16S539	D18S51
1	22,25	17,17	11,11	8,11	10,16	9,11	12,13	13,13
2	22,25	17,17	11,11	8,11	10,16	9,11	12,13	13,13
3	24,24	18,18	12,14	9,11	8,13	11,11	11,13	14,14
7	20,23	15,15	9,11	8,12	13,16	11,12	10,10	14,17
Sample	D19S433	D21S11	FGA	TH01	TPOX	vWA	CSF1PO	Amel
1	13,15	29,30	21,25	6,9	8,9	16,16	11,11	XY
2	13,15	29,30	21,25	6,9	8,9	16,16	11,11	XY
3	13,14	28,31.2	21,23	7,9	8,9	21,21	11,12	XY
4	12,14	28,32.2	24,24	6,8	8,11	16,16	8,11	XY

Probability of match

The example shown in Table 16.10 shows a complete match between the samples at serials 1 & 2. However, there is a remote probability that the match could be by chance. The probability of match by chances is inversely proportional to the number of loci tested. Its calculation in a given population requires comprehensive knowledge of the allele frequencies of all the loci tested. A worked example of the profile frequency (probability of match by chance) was discussed in Table 16.5.

Sexual assault

In a sexual assault the DNA testing is done to demonstrate male DNA from the perpetrator in a sample collected from the victim. In a vaginal swab the victim's own DNA is seen mixed with the DNA of the perpetrator. In addition the DNA from the husband may also be present. Less often there may be more than one perpetrator and in that case complex mixtures might be found. If the perpetrator has used condom his DNA would not be found in the vaginal secretions/swab. In that case the evidence material may have to be collected from other places like seminal stains on the victim's clothing or the other objects.

Example

The STR profiles of DNA extracted from high vaginal swab of a victim of sexual assault (sample 1) and the suspects (sample 2 & 3) are shown in Table 16.11. The vaginal swab shows mixture of two DNAs. The minor component has a male genotype and matches exactly with that of the suspect 2 whereas the suspect 3 is excluded.

Table 16.11 DNA profiles of victim (sample 1) and two suspects (samples 2 &3) in a case of sexual assault. The vaginal swab shows mixture of two DNAs. The minor component has a male genotype and matches exactly with that of the suspect 2 whereas the suspect 3 is excluded.

Sample	D3S1358	D5S818	D7S820	D8S1179	D13S317	D18S51
1	16,17(16,18)	11,13(10,11)	11,11(10,12)	11,14(13,15)	11,11(11,11)	12,18(13,15)
2	16,18	10,11	10,12	13,15	11,11	13,15
3	15,16	10,13	11,13	13,15	10,13	12,15
Sample	D21S11	TPOX	FGA	Th01	Amgl XY	-
1	29,30(29,32.2)	8,8(8,8)	23,24(22,23)	9.3,9.3(6,9)	XX(XY)	-
2	29,32.2	8,8	22,23	6,9	XY	-
3	28,31.2	8,8	21,22	6,9.3	XY	-

Resolving mixtures of DNA

Mixture of DNA from more than one source is typically encountered in investigation of sexual assault.

1. Simple mixtures may be resolved on gel electrophoresis.
2. Complex mixtures are best resolved on genetic analyzer.
3. In a mixture the major component is usually of the victim itself.
4. The minor component may be from one or more individuals.
5. Identify all possible alleles.
6. A difficult issue is to differentiate between stutter products and the minor component of the mixture.
    a. The stutter products are usually of smaller size than the main PCR product and therefore would be seen moving ahead of the main product.
    b. The height of the stutter peak is usually less than 10% of the main peak.
    c. The mixture allele peak when present in the stutter peak area is considered significant when its height is more than 15% of the main peak.
    d. An allele peak or band seen in the larger size area is unlikely to be a stutter peak/band.
    e. A peak height less than 10% present in an area where stutter peaks are not expected are considered significant.
7. The alleles once identified are sorted out.
8. The victim's alleles are identified by aligning/comparing them with her own DNA extracted from blood.
9. The DNA samples of the suspect(s) are run and the results are aligned/compared to find any match.
10. In a sexual assault male DNA can also be demonstrated by Y-STR profiling.

Paternity testing

At any autosomal locus an individual inherits one allele each from the biological parents (mother and father). The child inherits mitochondrial DNA only from the

mother while Y-chromosome is transmitted from the father to the son. The exceptions to the rules are development of spontaneous mutations or chromosomal aneuploidies (trisomy or monosomy). Most parentage disputes are of paternal in origin. However, occasionally maternity may also be questioned e.g. exchanged babies in a labour room.
In solving a paternity dispute a battery of STR markers are used. The step wise procedure includes:

1. Profiling of the alleged father and the child or the products of conception is done. The mother's profile is usually not required.
2. Of the two paternal alleles at each locus the allele transmitted to the child, called the "obligate allele", is selected.
3. If none of the alleles of the alleged father are present in the child at any of the loci tested the paternity is excluded. For example if the child has genotype 14,18 and the father has 13,15 the paternity is excluded. However, keeping the possibility of spontaneous mutations in mind it is advisable to consider more than one loci before excluding paternity.
4. Frequencies of the obligate paternal alleles are noted from the table of allele frequencies in the reference population.
5. The Paternity Index (PI) is calculated by dividing the prior probability with the frequency of the allele in question. The prior probability is the chance of transmitting the obligate allele by the alleged father to the child. If the alleged father is homozygous for the allele the prior probability is 1.0 and it is 0.5 if he is heterozygous.
6. Combined Paternity Index (CPI) is calculated by multiplying the PI values calculated at each locus.
7. The Probability of Paternity (POP) is calculated by the formula:
   CPI /CPI + (1-prior probability) x 100

Example

An example of calculation of paternity index and the combined paternity index is shown in Table 16.12.

Table 16.12 Calculation of Paternity Index and Probability of Paternity.

Locus	Genotype			Obligate Allele	Frequency	Paternity Index (PI) = Prior probability*/frequency
	Mother	Child	Al/Father			
D3S1358	14,15	14,18	15,18	18	0.078	0.5/0.078 =6.41
D5S818	9,12	9,9	9,9	9	0.049	1.0/0.049 =20.41
D7S820	8,11	8,9	9,11	9	0.108	0.5/0.108 =4.63
D8S1179	10,12	10,15	11,15	15	0.211	0.5/0.211 =2.37
D21S11	29,32.2	29,31.2	28,31.2	31.2	0.123	0.5/0.123 =4.06
TPOX	9,11	9,11	8,11	11	0.333	0.5/0.333 =1.50
TH01	8,9.3	9,9.3	6,9	9	0.250	0.5/0.250 =2.00
FGA	21,21	20,21	20,24	20	0.093	0.5/0.093 =5.38
Combined Paternity Index (CPI)			6.41 X 20.41 X 4.63 X 2.37 X 4.06 X 1.50 X 2.00 X 5.38 = 94,072 or 1 in 94,072			
Probability of Paternity			CPI/CPI + (1-prior probability) X 100 94072 /(94072 + 0.5) X 100 = 99.999 %			

*If the father is heterozygous for the allele Prior Probability = 0.5
*If the father is homozygous for the allele Prior Probability = 1.0

Identification in mass disasters

In mass disasters the DNA testing is primarily done to identify bodies that are beyond recognition. It may also be required to put the pieces of bodies together. The process is done as follows:

1. The DNA profiles of the dead bodies or their remains are entered in a computer database.
2. The DNA profiles of the bodies is matched with those of the parents or the offsprings. If these are not available then DNA may be matched with that recovered from the personal effects of the victims.
3. The matched profiles are put together after sorting.

Table. 16.13. Comparison of the DNA profiles of three dead bodies from a mass disaster and a father. The DNA profile of the "father" shares at least one allele at all fifteen loci with the DNA profile of the body number 7 and is therefore a proof of identity.

Sample	D2S1338	D3S1358	D5S818	D7S820	D8S1179	D13S317	D16S539	D18S51
3	20,25	14,15	8,10	8,12	14,16	10,11	10,10	13,16
7	20,23	15,15	9,11	8,12	13,16	11,12	10,10	14,17
16	22,23	14,14	10,11	8,11	13,16	10,12	10,12	15,17
Father	21,23	15,16	9,11	7,12	13,14	12,14	10,11	13,17
Sample	D19S433	D21S11	FGA	TH01	TPOX	vWA	CSF1PO	Amel
3	11,14	29,32.2	23,23	6,9.3	8,11	15,16	8,12	XY
7	12,14	28,32.2	24,24	6,8	8,11	16,16	8,11	XY
16	13,14	30,32.2	22,24	6,9	8,10	14,14	8,11	XY
Father	13,14	32.2,32.2	21,24	7,8	8,10	15,16	8,11	XY

Determining the ethnic origin of a person

Can DNA testing provide information on the ethnic origin of a person? The question is asked more than often especially when no clue is available to identify the perpetrator. There are some STR loci that have more significant differences in allele frequencies between major world populations like Caucasians, Blacks, Hispanics, Asians and East Asians etc. The differences are more marked in the frequency of less common alleles than the more common ones. An important requirement would be to have representative samples of the target ethnic groups with ethnically pure individuals and not the ones with self declared ethnicity. In this context analysis of SNP is considered more informative than the STR markers. In tribal populations with high frequency of consanguinity and marriage within the same tribe founder effect and genetic drift might cause an unexpectedly higher or lower frequency of alleles than in the rest of the population.

OmniPop 200.1 software is available on the internet free of cost. It is basically a database of the STR allele frequencies published in over 200 studies in the major world populations. By entering the CODIS STR profile of an individual it provides statistical information on finding a similar profile in various world populations. Similar software can be developed for one's own requirements if comprehensive knowledge of the allele frequencies in the smaller target groups and populations is available.

Collection storage and dispatch of samples for forensic DNA testing

1. Collection and storage of samples

a. DNA can be extracted from any biological material that contains sufficient number of nucleated cells. Since DNA is susceptible to degradation by mal handling and nucleases derived from cells and bacterial contamination the sample should remain as clean and free of bacterial contamination as is possible.
b. DNA may be extracted from a wide variety of samples. The usual samples are whole or dried blood (stains), buccal smear, wet or dried semen and other body secretions, hair with roots, soft tissues, fresh and dried bones etc.
c. The samples of stains like blood or body secretions should be air dried and kept in a paper envelop rather than plastic bag.
d. Fresh blood: 2-3 ml of venous blood should be collected in EDTA. The sample can be stored before dispatch at 4oC for 48 hours.
e. Blood stains on clothes or other objects should be dried at room temperature before dispatch in a paper bag.
f. Buccal smear obtained on a clean throat swab contains mucosal cells to yield sufficient quantity of good quality DNA. The swab can also be used to obtain DNA from a dead body. The swab should be rubbed several times over clean part of buccal mucosa. It may be air dried before storage or dispatch.
g. Semen and other body fluids containing nucleated cells are a good source of DNA. The stained clothes or objects should be air dried and treated as blood stains.
h. Hair shafts do not contain nuclear DNA. Hairs that are plucked from the body and come out with roots can be used to extract DNA. Sufficient DNA can be extracted from 2-3 hairs with roots.
i. Soft tissues are a very rich source of DNA. Skin is an easily accessible tissue that can be used to collect DNA from a dead body. A full thickness piece of skin measuring 2x2 cm from a clean part of the body or its remains can be taken. If skin is not available then any other available soft tissue should be collected. The soft tissue samples provide an excellent medium for bacterial growth. In a putrefied or heavily contaminated soft tissue sample the yield as well as the quality of DNA can be very poor. The soft tissue samples can be stored as such at -20oC in a suitable container for several days. The sample may be transported in normal saline. Do not put the samples for DNA testing in formalin at

any stage.

j. Bones can also be used as a source of DNA. But the extraction of DNA from a bone is difficult therefore the yield and quality of DNA is also variable. DNA is best collected from compact bones like humerus or femur. The spongy bones with thin cortex are usually heavily contaminated with mud etc. that may inhibit PCR.

2. Dispatch of samples

a. The samples should be properly labeled and sealed.
b. The request form should contain all available details of the individual(s) to be tested along with a brief summary of the incident and what exactly is required to be solved by the DNA test.
c. Unnecessary delay in transport can adversely affect the quality of the sample. There is no special requirement of transporting the samples in ice etc. However, avoid exposing the sample to extreme heat or direct sunlight.

3. Chain of custody

a. Forensic DNA testing is done for medico-legal purpose therefore it is essential to maintain the chain of custody.
b. A record of the individuals receiving and handing over the samples must be maintained at all steps as they may be called by the court as a witness.

Bibliography

1. Kaiser L and Sever G (1983) Paternity testing: I. Calculation of paternity indexes. Am J Med Genet 15: 323-329.
2. Elston RC (1986) Probability and paternity testing. Am J Hum Genet 39: 112-122.
3. Smith RN (1995) Accurate size comparison of short tandem repeat alleles amplified by PCR. Biotechniques 18: 122-128.
4. Brinkmann, B., Klintschar, M., Neuhuber, F., Huhne, J., and Rolf, B. (1998). Mutation rate in human microsatellites: Influence of the structure and length of the tandem repeat. Am J Hum Genet 62: 1408-1415.
5. Short Tandem Repeat DNA Internet DataBase, http://www.cstl.nist.gov/strbase/

17. Study of Donor Chimerism

The outcome of a stem cell transplant can be complete donor chimerism (1 0 0 % donor cells) or mixed donor chimerism of varying proportions of donor and recipient cells. Another type of chimerism called split chimerism may also exist in which one or more whole lineage is of recipient and the other of donor in origin.

The haematopoietic cells of donor origin can be detected in the host. The test may be done on peripheral blood, bone marrow or lineage specific cells e.g. T cells, B cells and granulocytes. Peripheral blood is equally sensitive in detection of chimerism than bone marrow. The study of donor chimerism may be done to know:

- Whether the donor engraftment has occurred or not?
- Whether there is mixed chimerism? If present then how much?
- If there is mixed chimerism then which lineages are mixed and which are fully donor?
- Whether there is chimerism in the lymphoid and the myeloid compartments?

Applications

- Myeloablative Stem Cell Transplant
  - Donor engraftment
  - Graft rejection
  - Prediction of GVHD
- Non-Myeloablative Stem Cell Transplant (Mini transplant)
- Relapse prediction

Techniques

Donor chimerism is usually tested by cytogenetics/FISH, real time PCR or STR analysis. STR analysis is basically an extension of the analysis of DNA mixtures discussed in Chapter 9 and 15.

Samples required

1. Recipient's pre-transplant blood sample in EDTA. If this is not available recipient's buccal mucosal cells, skin biopsy or hair roots may be collected on a stick swab to represent the pre-transplant status. Care should be taken to avoid contamination of the swab with recipient's blood.
2. Donor blood in EDTA.
3. Recipient's post transplant blood or bone marrow sample.

Procedure

1. Extract DNA from the three samples.
2. The STR profiling may be done by manual method or by genetic analyzer

(Chapter 9).

3. Out of the many STR markers "informative" marker(s) are chosen. The STR marker is called "informative" when its alleles can distinguish between the recipient and the donor DNA. For example if at the D21S11 locus the recipient has alleles 28,29 and the donor has 28,31 the marker is informative because allele 29 and 31 are exclusive for the recipient and the donor respectively.
4. In the presence of complete donor chimerism the recipient's post transplant DNA shows the donor's genotype. In mixed chimerism mixture of recipient and donor genotypes is seen (Fig 17.1).
5. In the manual method quantitative estimation of amplified products is done by densitometry of polyacrylamide gels (Fig 17.2).
6. If genetic analyzer is used for STR genotyping the analysis is done by measuring the peak heights and area under the curve.

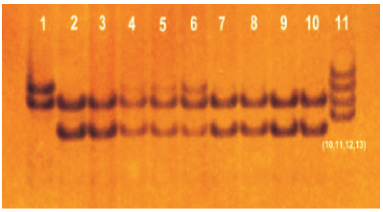


Fig. 17.1. PAGE of PCR amplification at D5S818 locus. Lane 11 shows allelic ladder with alleles 10-13. The lane 1 shows recipient's pre-transplant sample (alleles 11,12) and lane 2 shows donor sample (alleles 9,11). Lane 3-10 show serially collected recipient's post-transplant samples. Lanes 4-6 show gradually appearing recipient's exclusive allele (12) and a gradually decreasing strength of the donor' exclusive allele (9) indicating graft failure. The patient received an infusion of donor lymphocytes that resulted in disappearance of the recipient's allele (lanes 7-10).

Calculation of donor and recipient component

Calculation of the donor and recipient components is usually done by analyzing the samples on a genetic analyzer and then measuring the peak areas. A cost effective alternate is by doing PAGE followed by densitometry. Figs. 17.2 and 17.3 show an example of the calculation of donor chimerism by PAGE and its densitometry at the D5S818 locus.

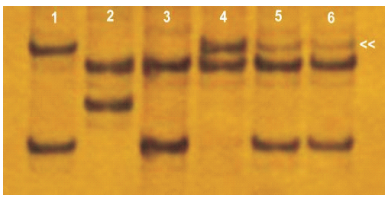


Fig. 17.2. PAGE of STR amplification at D5S818 locus. Lanes 1-4 show the samples of father, mother, donor and recipient (pre-transplant) respectively. Lanes 5 & 6 show the recipient's post-transplant samples in duplicate. The sample shows mixed donor chimerism represented by reappearance of the recipient's exclusive allele (arrow).

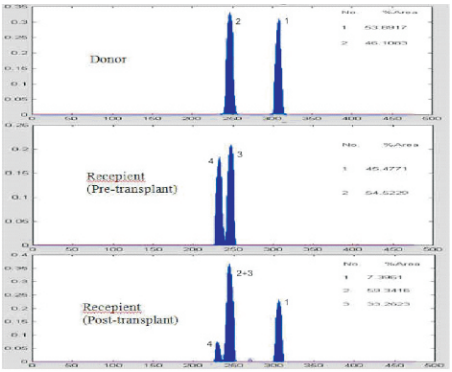


Fig. 17.3. Densitometric recording of STR amplified products run on polyacrylamide gel shown in Fig. 17.2. The recipient's post transplant sample shows mixed donor chimerism. The recipient's pre transplant allele peak 4 represents 7.3%. The total recipient's component is calculated by doubling the allele peak 4 component (3+4) i.e. 14.6%.

Real time PCR for assessing donor chimerism

As discussed in Chapter 6 end point analysis of PCR products by gel electrophoresis, including analysis on a genetic analyzer, is not very good for quantitative assessment of DNA. Qantitative assessment of donor chimerism by real time PCR can give more accurate measurement of the donor or recipient components. SNP genotyping by real time PCR using TaqMan® probe assay has been used for this purpose.

Bibliography

1. Scharf SJ, Smith AG, Hansen JA, McFarland C, Erlich HA (1995) Quantitative determination of bone marrow transplant engraftment using fluorescent polymerase chain reaction primers for human identity markers. Blood 85: 1954-1963.
2. Antin JH, Childs R, Filipovich AH, Giralt S, Mackinnon S, Spitzer T, Weisdorf D (2001) Establishment of complete and mixed donor chimerism after allogeneic lymphohematopoietic transplantation: Recommendations from a workshop at the 2001 tandem meetings. Biology of Blood and Marrow Transplantation 7: 473-485.
3. Thiede C, Lion T (2001) Quantitative analysis of chimerism after allogeneic stem cell transplantation using multiplex PCR amplification of short tandem repeat markers and fluorescence detection. Leukemia 15: 303–306.
4. Alizadeh M, Bernard M, Danic B, Dauriac C, Birebent B, Lapart C, Lamy T, Le Prise´ P, Beauplet A, Bories D, Semana G, Quelvennec E (2002) Quantitative assessment of hematopoietic chimerism after bone marrow transplantation by real-time quantitative polymerase chain reaction. Blood 99: 4618-4625.

Glossary of Terminology

Allele: one of two or more versions of a genetic sequence at a particular location in the genome.

Allele Frequency: proportion of an allele among the individuals in a population.

Allelic ladder: mixture of amplified DNA products of known sizes used for size comparison in gel electrophoresis of DNA.

Amplification: increase in the number of copies of a specific DNA fragment.
Amplification Refractory Mutation System (ARMS): application of PCR for detection of point mutations and SNP in which the DNA is amplified by allele specific primers.

Annealing: process of alignment/attachment of a short piece of DNA (primer or probe) to its complementary region on the target DNA.

Autosome: chromosome not involved in sex determination. The diploid human genome consists of 46 chromosomes including 22 pairs of autosomes, and one pair of sex chromosomes.

Base pair (bp): two nitrogenous bases paired together in double-stranded DNA by weak bonds (adenine with thymine and guanine with cytosine).

Base sequence: order of nucleotides in a DNA molecule.

BLAST: Basic Local Alignment Search Tool.

cDNA: DNA copy of RNA.

Chromosome: coiled up molecule of DNA with a supporting protein structure that carries the genetic information which is physically transmitted from one generation to the next.

CODIS (COmbined DNA Index System): database created by FBI that stores the DNA profiles of convicted offenders and the biological material found at crime scenes.

Coefficient of inbreeding: probability that an individual receives at a given locus two genes that are identical by descent (copies of a single gene carried by a common ancestor).

Combined Paternity Index (CPI): calculated by multiplying the PI values calculated at each locus.

Complementary sequences: DNA sequences that form a double-stranded structure by matching base pairs e.g. C-A-T-G is complementary to G-T-A-C.

Compound heterozygote: two different alleles at the same locus on the maternal and the paternal chromosomes (homologus chromosomes).

Cross-homology: two DNA (or protein) sequences are said to be homologous if they are sufficiently similar to suggest that they are derived from a common ancestral gene.

Ct (Cycle threshold): first significant increase in the amount of amplified product as

seen in the real time graphic recording of a PCR.

Cycle sequencing: technique in which the DNA to be sequenced is first amplified by a pair of primers. The amplified target is then subjected to linear or asymmetric amplification by a single primer i.e. "sequencing primer".

Degraded DNA: DNA broken into smaller fragments by shearing force or bacterial enzymes.

Denaturation: physical process of splitting the complementary strands of DNA to form single strands e.g. by heat or chemicals.

Deoxyribonucleic acid (DNA): genetic material of organisms, usually double-stranded; that contains deoxyribose, and four nucleobases.

DNA sequence: order of the base pairs in a DNA molecule.

DNAse: enzyme, usually of bacterial origin, that can degrade DNA into smaller pieces.

Donor chimerism: existence of two genetically different tissues in the same individual for example after receiving organ transplant.

ΔRn: difference between Rn+ and R-.

Electrophoresis: technique of separating charged molecules by electric current.

Entrez: an integrated, text-based search engine and retrieval system for the NCBI website.

Exon: portion of a gene that encodes amino acids.

FASTA format: text-based format for writing the nucleotide or peptide sequences using single letter codes.

Forward and reverse strands: two strands of a double-stranded DNA molecule each runs in opposite direction.

FRET Probe: a pair of probes that anneal to the inner region of the target DNA in head to tail configuration. Two florescent dyes are attached at the 3' and 5' ends of the upstream and the downstream probes respectively. On annealing the close proximity of the two fluorescent dyes ensures transfer of energy from the former to the later that in turn emits light of yet another wavelength.

Gene: an ordered sequence of nucleotides located at particular position on particular chromosome and encodes specific functional product like protein or RNA.

Gel: semisolid matrix used in electrophoresis to separate molecules e.g. agarose and acrylamide etc.

GenBank: an annotated collection of all publicly available DNA sequences in the form of database at the National Institute of Health (NIH), Bethesda USA.

Gene frequency: relative frequency of a particular allele in a population.

Genome: total genetic material found in a cell; its size is generally given as number of base pairs.

Genotype: an organism's complete collection of genes. The term also refers to the two alleles inherited for a particular gene.

Heterozygote: different alleles at the same locus on the maternal and the paternal chromosomes (homologous chromosomes).

Homozygote: similar alleles at the same locus on the maternal and the paternal chromosomes (homologous chromosomes).

Intron: intervening sequence of DNA between two coding regions (exons).

Linkage: proximity of a marker and a gene on a chromosome; when the marker and the linked gene are close together the probability of their separation during meiotic crossover is low and the two are inherited together as a unit.

Locus: specific physical location on a chromosome.

Mini STRs: STR analysis in a degraded DNA sample can be improved by bringing the forward and the reverse amplification primers flanking the repeat units to the nearest possible distance from the repeat units.

Mitochondrial DNA: each mitochondrion contains 2-10 copies of a circular piece of DNA, 16,569 base pairs in length. It is inherited from the mother.

Molecular Beacon probe: short sequence of DNA complementary to the internal region of the PCR target DNA. The un-annealed probe assumes the shape of a hairpin due to the presence of short DNA sequences complementary to each other. The later also ensures quenching of the attached fluorescent dye and the quencher. On annealing the hairpin structure opens up relieving the fluorescent dye from the effect of quencher.

Multiplex PCR: PCR done with more than two primers.

Mutation: heritable change in DNA sequence that is associated with a harmful effect on the phenotype.

Next generation sequencing (NGS): new technique of extremely high throughput genomic sequencing involving template preparation, sequencing, imaging, and data analysis. The template preparation is done by emulsion PCR or solid phase amplification. The templates are immobilized on a solid surface on which thousands to billions of separate sequencing reactions can take place. The sequencing is done by fluorescent labeled nucleotides. The data are aligned with a known reference sequence.

Nucleotide: basic building block of DNA comprising phosphate, ribose or deoxyribose, and a purine or pyrimidine base.

Obligate allele: of the two parental alleles at each locus the allele transmitted to a child.

Oligo dT: primers that are designed to anneal with the poly-A tail attached at the 3' end of mRNA molecules.

Oligonucleotide: short, typically synthetic, polymer of nucleotides.

Oncogene: gene which is associated with cancer.

Paternity Index (PI): calculated by dividing the prior probability with the frequency of the allele in question.

Phenotype: observable traits of an individual such as height, eye color, and blood group etc.

Point mutation: alteration in DNA sequence caused by a single-nucleotide base change, insertion, or deletion.

Polymerase chain reaction (PCR): in-vitro technique in which a DNA molecule can be amplified in sequential steps of denaturation, primer annealing and extension to produce millions of copies of desired length.

Polymorphism: inheritable change in DNA sequence that is present in more than 1% individuals of a population. Unlike mutation polymorphism is not associated with any harmful effect on the phenotype.

Primer: short sequence of single-stranded DNA of a specific base composition to which new deoxyribonucleotides can be added by DNA polymerase.

Prior probability: probability of transmitting the obligate allele to the offspring. It is 1.0 (100%) when the parent is homozygous and 0.5 (50%) when the parent is heterozygous for an allele.

Probability of Paternity (POP): is calculated by the formula: CPI/CPI + (1-prior probability) x 100.

Probe: short sequence of single-stranded DNA or RNA of a specific base composition labeled with radioactive, chemi-lumiscent or fluorescent dye.

Proteinase-K: broad spectrum protein cleaving enzyme derived from extracts of fungus Engyodontium album and commonly used in nucleic acid extractions.

Quencher: compound that has the ability to suppress the light emitting from a fluorescent dye. The quencher may be a high energy fluorescent dye like rhodamine (TAMRA) or a non-fluorescent chemical quencher e.g. DABCYL and black hole quenchers (BHQ).

Random hexamers: mixture of all possible combinations of six nucleotide pieces of DNA.

Real time PCR: in-vitro technique of DNA amplification in which the process is monitored by evaluating amplification at the end of every cycle. The amplification is monitored by fluorescent dyes.

Restriction enzyme: enzyme of bacterial origin that can cut DNA at pre-defined sequences.

Restriction fragment length polymorphism (RFLP): variation between individuals in DNA fragment sizes cut by specific restriction enzymes.

Reverse transcriptase: enzyme used for synthesizing DNA copy of RNA.

Rn: indicator of the amount of fluorescence generated in a PCR. Rn of a test reaction and a non target control (NTC) are called Rn+ and Rn- respectively.

RNase: enzyme that can degrade RNA into smaller pieces. It is present in all

organisms.

Sequencing: in-vitro technique to determine the order of nucleotides in DNA or RNA molecule or amino acids in a protein.

Short tandem repeats (STR): tandemly repeated short stretches (units) of DNA (2-6bp) that are randomly distributed in the genome. The number of repeat units vary between individuals of a population.

SNP (single nucleotide polymorphism): naturally occurring variations of DNA present after every 300-500 nucleotides in the genome.

Stutter bands: in PCR amplification of STRs the elongating DNA strand can slip one or more steps forward resulting in formation of end products that are one or more repeat units smaller in length.

SYBR green: dye that binds to minor groove of double stranded DNA. It gives little fluorescence when in solution but emits a strong fluorescent signal after binding with double-stranded DNA. It can be used to monitor real time PCR.

Tm: temperature at which 50% of the primer/probe is annealed to its target.

TaqMan® probe: short sequence of DNA complementary to the internal region of the PCR target DNA. The 5' end of the probe is labeled with a fluorescent dye (reporter) while its 3' end is labeled with a fluorescence quencher (suppressor). It is used for real time monitoring of PCR.

Trisomy: presence of three homologus chromosomes instead of the usual two.

Twin peaks "+A" and "–A" products: Taq polymerase has the property of adding +A at the 3' end of the elongating strand that results in formation of products one base pair longer. These products appear as twin peaks on a genetic analyzer.

VNTR: Variable Number of Tandem Repeats are tandemly repeated sequences of DNA that range in length between 7-80bp. These are also called mini-satellites.

Index

About the author

Maj Gen Suhaib Ahmed, HI (M) graduated from King Edward Medical College in 1980. He is a fellow of the College of Physicians and Surgeons, Pakistan since 1986. He also did PhD from University College London in 1998. While working at the prestigious Armed Forces Institute of Pathology, Rawalpindi he has the honour of starting in 1994 the first clinical service for prenatal diagnosis of thalassaemia in Pakistan. Since then he has done over 7000 prenatal diagnoses. He is renowned for developing several cost effective PCRs in neoplastic and infectious diseases. He also has plenty of experience in using STR analysis for diagnostic and forensic use. He has over 50 publications in the National and some of the most well reputed International Journals.

ISBN - 978-969-9829-00-0