Origins of . . .

The polymerase chain reaction in pathology

J J O’Leary, K Engels, M A Dada

A surprisingly simple method of making unlimited copies of DNA fragments conceived under unlikely circumstances—during a moonlit drive through the mountains of California.

Kary B Mullis

The movie Jurassic Park and real life events like the O J Simpson trial have captured public interest in molecular biology. The basis of molecular biology is the understanding of the structure and function of DNA, RNA, and proteins, as well as the techniques for manipulating these molecules. In this review we demonstrate how a single technique revolutionised this research area and made molecular biological methods accessible, not only to research groups but also to diagnostic pathology laboratories. As a result, molecular pathology is now firmly established as a pathology discipline, providing new insights into the pathogenesis of disease as well as innovative techniques in diagnosis.

Historical background

Although DNA was first isolated in 1869 by Miescher, its double helix structure was not described until 1953 by Watson and Crick. In 1955, Arthur Kornberg of Stanford University discovered DNA polymerase. This cellular enzyme is involved in DNA replication and repair by catalysing the addition of nucleotides to the 3' end of an existing DNA chain. The initiation of a new chain requires an existing oligonucleotide or polynucleotide chain, referred to as a primer. The polymerase attaches nucleotides in a new DNA strand, complementary to nucleotides on corresponding positions of the parent DNA strand (template strand). DNA polymerases are involved in the assembly of RNA from a DNA template (transcription).

Over time, new tools for producing and manipulating DNA were developed. Restriction endonucleases (RE) cut DNA at specific sequences (restriction sites), making it possible to isolate strands of DNA containing specific genes. In 1975, Edwin Southern described a technique for the localisation of specific sequences within genomic DNA by electrophoretic transfer techniques. This technique, subsequently known as Southern blotting, involves the digestion of genomic DNA by one or more REs and the separation of the resulting fragments by agarose gel electrophoresis. The separated fragments of double stranded DNA are then separated into single stranded form and transferred (blotting) from the gel to a solid support (usually a nitrocellulose or nylon filter). The sequence of interest can then be detected using a short fragment of DNA (oligonucleotide probe), which is complementary to the DNA sequence of interest (hybridisation).

Initially, radioactive labels (\(^{32}\text{P}, {35}\text{S}, \text{H}\)) were used for probing, but later non-isotopic labels including biotin, digoxigenin, and fluorescein were employed. Three nucleic acid labelling methods are now described, including enzyme incorporation, chemical derivatisation, and chemical cross linking. The enzymatic labelling reactions for DNA include nick translation, random priming, and 5' and 3' end labelling using DNA polymerase, Klenow polymerase, and Dnase 1. For RNA detection, riboprobes can be created using SP6, T3, and T7 in vitro enzymologies. Alternatively, synthetic oligonucleotides can be conveniently used as probes for DNA and RNA detection assays. Additionally, PCR probe labelling methods, employing biotin, digoxigenin or fluorescein labelled dNTPs in substituted molar ratios in the PCR reaction mix, can also be used for generating probes for any hybridisation analysis.

A refinement of Southern blot analysis, of interest to pathologists, was first published in 1969 by two groups working independently in the UK and the USA. The technique—in situ hybridisation—allowed for the first time direct correlation between hybridisation signals and tissue morphology. The initial reports were followed by application of the methodology to cryostat, paraffin wax, chromosomal, and electron microscopy preparations. Since then, numerous DNA and RNA targets have been demonstrated using both isotopic and non-isotopic labels.

Analogous techniques for RNA and proteins have been named northern and western blotting, respectively, as a play on the name Southern. These techniques gained widespread acceptance in the 1970s.

The next significant breakthrough in molecular biology was the development of rapid DNA sequencing techniques. In the chemical method of Maxam and Gilbert the sequence is determined from native DNA itself. DNA is labelled at one end and then exposed to agents that destroy one or two of the nucleotides resulting in fragments that can be separated and analysed by electrophoresis. The Sanger or dideoxy chain termination method parallels the
process of DNA replication. Starting with a primer, a DNA polymerase adds nucleotide triphosphates (dNTPs) producing a complementary DNA strand. Four different reaction mixtures, each containing in addition a deoxyribonucleotide triphosphate (dNTP) corresponding to one of the four nucleotides, is added to the new DNA strand and terminates the replication process. This produces DNA fragments of various lengths. The sequence is then determined by separating the resulting DNA fragments of each reaction by gel electrophoresis.

For further progress to be made, techniques allowing the production of large quantities of recombinant DNA were necessary. Cloning was the first revolutionary technique described, involving the isolation and production of many copies of a DNA sequence. The DNA fragment of interest is cut using REUs and then ligated into other DNA molecules called vectors or cloning vehicles (for example, plasmids). The vector and the inserted DNA fragment can then be produced in large quantities by transformed bacteria. Subsequently, the cloned sequence can be extracted and analysed or used directly as a probe for hybridization.

Cloning is a time consuming process and is not routinely applicable to a busy diagnostic pathology laboratory. This obvious disadvantage was overcome by the development of the polymerase chain reaction (PCR). The technique was first described by Khorana and colleagues in the early 1970s, but brought to life and named PCR in 1983 by Kary Mullis, who subsequently received the Nobel Prize for Chemistry in 1994 for his work on PCR.

For developing a new method of DNA replication, Mullis initially envisaged the hybridisation of oligonucleotide primers to single stranded DNA and subsequent extension of these primers by a DNA polymerase, akin to the process of in vivo DNA replication in mammalian cells. After much variation and optimisation of the technique he was able to amplify a 25 base pair fragment of a plasmid using two oligonucleotide primers of 11 and 13 bases long. After this success, the first paper on PCR was published in Science in 1985 by seven scientists (including Mullis) from Cetus, a Californian biotechnology company.

The initial PCR method used the Klenow fragment of *Escherichia coli* DNA polymerase I to extend the annealed primers. As this enzyme is inactivated by the high temperatures used to melt (denature) DNA strands, it had to be replenished during every cycle after each denaturation step. With the discovery of thermostable DNA polymerases such as Taq (*Thermus aquaticus*) polymerase, the PCR process became simpler, obviating the need for fresh enzyme addition after each heating step. In addition, these enzymes are active at higher temperature, thus increasing specificity and the rate of DNA synthesis.

The subsequent automation of the PCR process using dedicated DNA thermal cyclers and its simplicity and ease of use led to its widespread application in disparate scientific disciplines such as cell biology, and medical specialties such as forensic medicine and tumour pathology. The impact of this development on research was acknowledged by honouring PCR with the title “Major scientific development of 1989” and Taq DNA polymerase “Molecule of the year 1989” by Science. Since then, there has been an explosion in the number of publications dealing with PCR applications.

Owing to financial impact, PCR technology has been the subject of ongoing litigation. Cetus, the US biotechnology company was originally granted the patent to native Taq DNA polymerase in 1989. Around the same time as the 1989 Science article, DuPont challenged the PCR patent. These patents were upheld in February 1991 and by the end of the year Cetus had disappeared and the Swiss pharmaceutical company Hoffmann-La Roche had acquired the patent rights for native Taq, recombinant Taq polymerase, and the PCR methodology for US$800 million. The legal drama continues with Roche now fighting the US laboratory supply company, Promega, over patent rights.

**Raw materials**

Many different types of clinical samples such as blood, semen, saliva, single hairs, archival fixed paraffin, and plastic embedded tissues can be used for DNA and RNA amplification. For Southern blot DNA detection, large amounts of DNA from most tissues may not be available. Indeed, archival paraffin wax embedded material up to 40 years old has been successfully used for DNA amplification, using PCR technology.

Tissues for histopathological examination are usually fixed in a suitable fixative to maintain morphology. The commonly used formaldehyde fixatives nick DNA and thereby reduce the maximal size of product that can be amplified. Good yields of nucleic acid can however be obtained using proteolytic enzymes such as proteinase K or pepsin. The alcohol based fixatives (Carnoy’s, methanol, methanol:acetic acid) also accommodate PCR amplification to greater or lesser degrees, while mercuric chloride based fixatives largely inhibit PCR.

The age of source material for PCR appears limitless. Even palaeobiological plant matter, up to 20 million years old has been used. DNA and RNA can be extracted from specimens by a variety of well described techniques. The precise technique depends on the type and amount of starting material (fresh or fixed), the amount of potential PCR inhibitors present (in the sample and type of nucleic acid being extracted). PCR inhibitors are ubiquitous and include potassium ions, porphyrins from haeme, and other undefined products. Most extraction methodologies employ phenol-chloroform extraction, which in most cases eliminates these inhibitors. Many proprietary DNA and RNA purification kits are now available, obviating the need for complex and time consuming extraction protocols.
Origins of... The polymerase chain reaction in pathology

Technique

The originally described PCR technique involved separating double stranded DNA and hybridising oligonucleotide primers (usually 17–30 nucleotides in length) to the different strands flanking the DNA sequence to be amplified. New DNA template was created by thermocycling the reaction, through denaturation, annealing, and extension phases (fig 1).

Various adaptations of PCR have been developed, many of which are now used in diagnostic pathology laboratories worldwide.

REVERSE TRANSCRIPTASE PCR
Reverse transcriptase PCR (RT-PCR) is used for detection of RNA targets. In this reaction, copy DNA (cDNA) is created using a reverse transcriptase enzyme (for example, MMV1V RT) and then subsequent amplification of the newly created cDNA follows. Originally the method used a two step procedure: reverse transcription and DNA amplification. The development of rTth polymerase, which combines reverse transcriptase and DNA polymerase activity, obviates the need for a two step reaction. This is a major improvement, as it minimises handling and lowers possible contamination risks.32–35

ASYMMETRIC PCR
Asymmetric PCR is a simple and effective method for the production of single stranded DNA suitable for direct sequencing. It uses unequal molar concentrations of primers in the reaction set up, essentially driving the reaction to single target strand accumulation. This can easily be sequenced directly, without the need for cloning or the establishment of DNA libraries.36

INVERSE PCR
Inverse PCR allows amplification of DNA outside the boundaries of known sequences. This is important in the study of viral tumorigenesis, when attempting to identify possible insertion sites of viruses in host DNA, and for the assessment of clonality in lymphoid tumours.37

AMPLIFICATION REFRACTORY MUTATION SYSTEM
Amplification refractory mutation system (ARMS) is a novel system using primers designed so that the 3’ end coincides with a mutated nucleotide base, facilitating allelic discrimination.38

SINGLE STRAND CONFORMATION POLYMORPHISM
Single strand conformation polymorphism (SSCP) is another method used for the detection of single base changes in DNA and RNA, and relies on the different mobilities of DNA strands containing single base pair differences, when run on non-denaturing polyacrylamide gels.39–42

DIFFERENTIAL DISPLAY
Differential display allows the simultaneous genetic analysis of changes in gene expression in cells and tissues. The technique uses a set of primers, one of which will hybridise to a polyadenylated tail present in mRNA (the primer also contains a one or two base anchor), the other primer is short and arbitrary in sequence, and anneals in different positions relative to the first primer. A combination of nearly 300 primers is required to ensure that each possible mRNA is amplified at least once. The mRNA populations defined by these primers are amplified after reverse transcription and resolved on a DNA sequencing gel. Fragments that display differential expression between the diseased and non-diseased states can easily be excised from the gel. Subsequent cloning of individual mRNAs is then possible. The technique has been particularly useful for the detection of differentially expressed genes in leukaemia, heart disease, and diabetes mellitus.40

CDNA SUBTRACTION PCR
CDNA subtraction PCR can be applied easily to cells enriched by flow cytometry. The subtraction procedure involves three steps leading to the identification of a collection of full length cDNAs cloned in an expression vector, suitable for direct functional analysis. In
the first step, an RT-PCR is performed that amplifies cDNA representing all poly mRNA present in two different samples (X and Y). The PCR produces 3’ cDNA stubs of approximately 200-600 base pairs that can be amplified through multiple rounds of PCR while maintaining the gene expression profile present in the starting mRNAs. The second step involves the reciprocal removal of common sequences from both samples using a biotin-avidin cDNA subtraction protocol. Subtraction product X - Y is enriched for sequences present in X but not in Y and similarly Y - X is enriched for sequences in Y, not found in X. The final step of the reaction involves labelling of the subtraction products X - Y and Y - X, which are used to screen replica filters from a full length library. cDNA clones are selected that hybridise consistently with one and not the other subtracted probe.

REPRESENTATIONAL DIFFERENCE ANALYSIS

Representational difference analysis is a new technique that allows the identification of the differences between two complex genomes. The technique basically involves a genomic subtractive hybridisation protocol, which allows the investigator to discriminate sequences present in a tumour specimen form normal control DNA of the same individual.

IN SITU PCR

In situ PCR is, for most histopathologists, the marriage of standard histopathology and molecular biology. In situ PCR is used to detect single copy target nucleic acid sequences in fixed tissues and cells. It aims to correlate PCR results with morphology. While holding the greatest potential for diagnostic histopathology, it is a technique that needs to gain widespread acceptance.

TAQMAN PCR

TaqMan PCR (5’ nuclease assay) is a major advance in PCR. It was first described by Holland et al., who used the 5’-3’ endonucleolytic activity of Taq DNA polymerase to detect target sequences during amplification by PCR. Included in the PCR mixture is a probe (usually 20-30 mers in length) designed to hybridise within the target sequence and to be non-extendible at the 3’ end. The fluorescent emission activity of a fluorescent reporter molecule attached to the probe at its 5’ end is neutralised by a quencher molecule at the 3’ end. When hybridised to its target sequence, the intact probe shows no signal because of the proximity of the reporter molecule to the quencher molecule. During amplification Taq DNA polymerase, through its 5’-3’ endonucleolytic activity, cleaves the probe into fragments, separating the reporter molecule from the quencher, thus allowing its detection. The amount of fluorescence is directly proportional to the amount of specific amplification of the target. The major advantage of this technique is its ability to detect specifically amplified DNA or RNA sequences at selected time points in the PCR, thereby allowing direct quantitative real time DNA and RNA detection. This is achieved using specifically designed equipment (for example, Perkin Elmer Applied Biosystems 7700 DNA sequence detector). Alternatively, an end point format can be adopted, in this case using a luminescence spectrometer.

COMPARATIVE GENOME HYBRIDISATION

Comparative genome hybridisation (CGH) is a new approach in fluorescence in situ hybridisation, allowing the comprehensive analysis of chromosomal imbalances in entire genomes. Genomic DNA from cell populations to be tested is labelled with modified nucleotides (dig 11dUTP) and used as a probe to normal metaphase chromosomes of the patient. This probe is called the test probe. As an internal control, genomic DNA derived from cells with a normal karyotype is differentially labelled (control DNA probe) and hybridised simultaneously with the test probe. For detection of the hybridised test, and control DNA probes, different fluorochromes are used and each is visualised with epi-fluorescence microscopy with selective filters. If the tissue under analysis contains additional chromosomal material, hybridisation reveals higher signal intensities at the corresponding target regions of the hybridised chromosome. Conversely, deletions are visible as lower signal intensities. By comparing the hybridisation patterns of the test and control probes, changes in signal intensities caused by allelic imbalance can be conveniently identified.

Applications of PCR in pathology

PCR is an established technique and has increased the range and sensitivity of diagnostic procedures. The exquisite sensitivity of PCR is also its major drawback, as contamination and amplification artefacts can give rise to difficulties in the interpretation of results.

Table 1 Microorganisms that can be detected by PCR

<table>
<thead>
<tr>
<th>Microorganisms that can be detected by PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viruses</td>
</tr>
<tr>
<td>Adenovirus</td>
</tr>
<tr>
<td>Cytomegalovirus</td>
</tr>
<tr>
<td>Epstein-Barr virus</td>
</tr>
<tr>
<td>Hepatitis A, B, C</td>
</tr>
<tr>
<td>Herpes simplex virus</td>
</tr>
<tr>
<td>HIV I and II</td>
</tr>
<tr>
<td>Human herpesvirus 7, 8</td>
</tr>
<tr>
<td>Human papillomavirus</td>
</tr>
<tr>
<td>HTLV-1</td>
</tr>
<tr>
<td>Lassa virus</td>
</tr>
<tr>
<td>Measles virus</td>
</tr>
<tr>
<td>Rotavirus</td>
</tr>
<tr>
<td>Bacteria</td>
</tr>
<tr>
<td>Mycobacterium tuberculosis</td>
</tr>
<tr>
<td>Mycobacterium tuberculosis</td>
</tr>
<tr>
<td>Mycobacterium leprae</td>
</tr>
<tr>
<td>Borrelia burgdorferi</td>
</tr>
<tr>
<td>Legionella pneumophila</td>
</tr>
<tr>
<td>Listeria monocytogenes</td>
</tr>
<tr>
<td>Chlamydia trachomatis</td>
</tr>
<tr>
<td>Helicobacter pylori</td>
</tr>
<tr>
<td>Protozoa</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
</tr>
<tr>
<td>Plasmodium falciparum</td>
</tr>
</tbody>
</table>

MICROBIOLOGY

In the past, diagnosis of infections was limited by the supply of appropriate material for culture, protein analysis or microscopy. These limitations have been overcome by the introduction of PCR in diagnostic microbiology. It is now possible to detect DNA or RNA of infectious organisms that are either present in small numbers, slow growing (viruses, mycobacteria, etc) or in material not suitable for culture. PCR can facilitate the diagnosis of early and latent stages of infection, which cannot be identified by conventional laboratory techniques.

The examination of archival material allowing retrospective studies has had great impact and has demonstrated correlations between viral agents and tumorigenesis (for example, human papillomavirus and cervical carcinoma, Epstein-Barr virus and post-transplant lymphoproliferative disorder (PTLD), and Kaposi sarcoma herpesvirus/human herpesvirus 8 and Kaposi’s sarcoma). Table 1 lists microorganisms detectable by PCR in routine clinical samples such as blood, cerebrospinal fluid, semen, saliva, faeces, pleural fluid, and fixed tissues.
HUMAN GENETICS
A major use of PCR is in the diagnosis of chromosomal disorders or hereditary diseases, such as Down's syndrome, β thalassaemia, cystic fibrosis, and haemophilia (table 2). Invasive antenatal procedures to obtain fetal cells, such as chorionic biopsies and amniotic fluid sampling, have an inherent risk to the fetus, and can perhaps be replaced by non-invasive techniques. Fetal DNA may be amplified from maternal blood by PCR, and fetal blood cells from maternal blood can be used for aneuploidy detection and to determine fetal sex.41-49 Fetal cells isolated from maternal cervical mucus have also been used for genetic analysis.46 Parental testing for genetic disease is made easier by PCR to detect variable numbers of tandem repeats (VNTRs), microsatellite tandem repeats, and allele specific sequences in the parental genome. This can be achieved using only a few cells with a fluorescent multiplex PCR approach, analysing “microsatellite fingerprints” and disease loci in one reaction (B Tutschek, personal communication).

TUMOUR BIOLOGY/ONCOLOGY
In oncopathology, PCR has led to a better understanding of the pathobiology of malignancy, allowing the analysis of mutations in oncogenes and tumour suppressor genes (for example, c-myc, p53, ras), the detection of minimal residual disease (MRD), clonality (for example, B and T cell gene rearrangements in lymphomas) in identifying gene rearrangements (for example, t(14,18) in follicular lymphomas and the Philadelphia chromosome in CML), and in the assessment of loss of heterozygosity (allelic imbalance) particularly in colorectal and breast cancer. PCR's greatest versatility is that it allows the examination of formalin fixed paraffin wax embedded tissue, in which DNA may be degraded and is therefore not suitable for Southern blotting. Using such archival material, large scale retrospective genetic analysis of p53, DCC, APC, and ras mutations in colorectal cancer, and genome-wide screening for novel tumour suppressor genes and oncogenes in any cancer can be easily undertaken using PCR and fluorescent amplicon detection technologies.51-55

FORENSIC PATHOLOGY
PCR has brought significant progress in forensic pathology. It is used in establishing the identity of mutilated corpses or decomposed human remains, in sex determination, in cases of disputed paternity, and in identifying perpetrators of crime. This is based on the amplification of VNTRs or restriction fragment length polymorphisms (RFLPs) and is referred to as DNA fingerprinting. PCR has not only made such evaluations easier, but also possible even with trace amounts or partially degraded biological material such as blood and semen stains or hair, increasing dramatically the range of samples that can be analysed.54 55

Where to from here?
PCR is a highly specialised research tool with many uses in medical laboratories. PCR methodology is well established, which greatly facilitates genetic, microbiological, and virological analysis. The advent of automated thermal cyclers, fluorescent DNA sequencers, and real time PCR sequence detectors (ABI 7700 DNA sequence detector) has also extended the power and repertoire of PCR.

The major advance of PCR is that it can amplify a sequence of DNA from among the background of the entire genome (three billion base pairs in a haploid cell), making it exquisitely more sensitive than other molecular biological tools.

In the past few years many startling advances have been made in PCR technology including in situ PCR, direct sequencing of PCR products, and quantitative assays using ELISA technology and real time PCR detectors, which automatically measure DNA and RNA loads directly in the starting sample.

New enzymology including the combined reverse transcriptase DNA polymerase enzyme (RTth DNA polymerase), the recently introduced long amplifying Taqs (such as TaqXL), and newer sequencing Taqs (such as Taq CS) will greatly facilitate the investigation of human disease, making the basic technique of PCR more robust and more easily reproducible.

Microchip based PCR technologies may have once seemed a dream, but we expect they will become reality in the next 5–10 years. It is envisaged that this will utilise a compartmentalised solid support with bound reporter probes, or acceptor molecules with the potential to identify point mutations, subchromosomal, genomic, or DNA sequence polymorphisms. This technique, which is non-invasive, will greatly simplify the diagnosis of human disease.

A simple idea may have far reaching implications; in Kary Mullis's own words, the simple method of making unlimited copies of DNA through PCR and application to clinical practice could be used to completely revolutionize our approach to the diagnosis and management of disease.

34 Gillenston UB, Erlich HA. Generation of single stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. Proc Natl Acad Sci USA 1988;85:7452-6.
54 Grillingston UB, Erlich HA. Generation of single stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA-DQA locus. Proc Natl Acad Sci USA 1988;85:7452-6.