Internal Medicine Grand Rounds

DNA Amplification by the Polymerase Chain Reaction: Applications to Clinical Medicine and Clinical Investigation

Joel D. Taurog, M. D.

August 31, 1989

I. INTRODUCTION AND BACKGROUND

A. Brief history of the polymerase chain reaction.

Late in 1985, a group of investigators from the Cetus Corporation reported a method for the enzymatic amplification of specific sequences of DNA using repeated cycles of denaturation of the DNA followed by hybridization of oligonucleotide primers complementary to the target DNA sequence and then extension of the primers with DNA polymerase [1]. In their report, they demonstrated 220,000-fold amplification of the ß-globin gene from small samples of human genomic DNA. Although the principles on which this method was based were not new, it represented a novel application of existing technology that greatly expanded the speed and sensitivity with which defined sequences of DNA could be cloned. Several reports during the following two years documented the power of this method, which was soon termed the polymerase chain reaction (PCR), although the tedium of the method as originally described undoubtedly limited its widespread use.

Early in 1988, the same investigators reported a significant improvement in the method that simultaneously simplified its performance and improved the sensitivity and specificity of the enzymatic reaction [2]. At about the same time, instruments for automating the procedure became commercially available. The availability of both the simplified enzymatic reaction and its automation, together with the inherent power of the method, has led to an explosion of reports in which PCR has been applied to a host of biomedical problems (Figure 1).

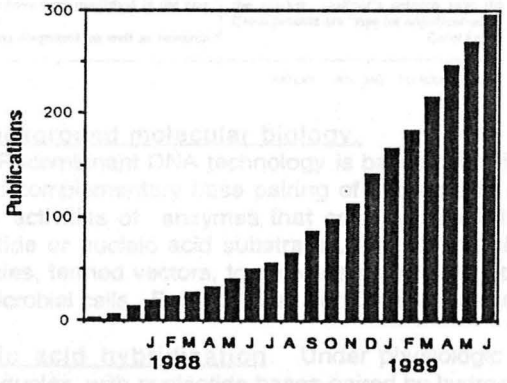


Figure 1. Cumulative publications listed under Polymerase Chain Reaction in BRS Colleague data base, by date of publication from reference 1 through June 1989. The first three bars represent 1985, 1986, and 1987.

In this Grand Rounds, I will present the principles underlying PCR and examples of its present and potential future applications to clinical medicine. I will not formally address financial considerations, either in regard to the proprietary claims of the Cetus Corporation and its competitors (Figure 2), or in regard to the expense of the procedure itself, since these are of little significance to the scientific importance of the method.

Rival claims over DNA amplification

Washington

DuPont is challenging Cetus Corporation's stranglehold on the lucrative market for the DNA amplification technique named the polymerase chain reaction (PCR). Last week, DuPont's attorneys asked a judge to rule that Cetus's PCR patents are invalid, but they will not disclose the basis for their argument.

Cetus received two key patents covering PCR just over two years ago. The PCR technique involves using two specific DNA probes that flank an unknown region of DNA to be studied. Through alternating cycles of heat and cooling, the probes squeeze between the strands of the unknown DNA and initiate an exponential copying process that can yield millions of copies of the unknown DNA sequence (see *Nature* 331, 461; 1988).

PCR has become widely used, and hundreds of research papers using the technique have been published in the past two years.

PCR has diagnostic as well as research applications: it has been used to screen fetuses for genetic disorders, and several companies are developing tests using PCR to detect genetic material from the AIDS virus. The market for diagnostic PCR kits could reach $1,500 million by 1998.

In a joint venture with Perkin-Elmer, Cetus sells a thermal cycling instrument to automate the PCR process. The cycler comes with two PCR reagent kits and costs $8,500.

Permission to use PCR in basic research is included in the cost of the cycler and reagents. But Cetus has licensed all diagnostic uses of PCR to Hoffman LaRoche, and has said it will prosecute those who use PCR without buying one of its machines.

DuPont is also developing 'nucleic acid replication kits' and a thermocycler, and plans to begin selling them later this year.

Now that its products are approaching the market, DuPont's petition says the Cetus patents are "ripe for adjudication".

Carol Ezzell

NATURE · VOL 340 · 10 AUGUST 1989

Figure 2. PCR and litigation.

B. Background molecular biology.

Recombinant DNA technology is based primarily on (a) the thermodynamically favored complementary base pairing of nucleic acid chains to form stable duplexes, (b) the activities of enzymes that catalyze several classes of reactions on either nucleotide or nucleic acid substrates, and (c) the ability of certain classes of DNA molecules, termed vectors, to replicate themselves upon introduction into appropriate host microbial cells. Each of these principles will be considered briefly in turn.

Nucleic acid hybridization. Under physiologic conditions, DNA consists of a helical duplex, with nucleotide bases paired by hydrogen bonding and the two strands running antiparallel, one 5' phosphodiester bond to 3' phosphodiester bond (5' to 3'), and the other 3' to 5', as shown in Figure 3. As the temperature is raised, the hydrogen bonding between the two strands is insufficient to offset the increased thermal energy, and the two strands begin to separate. For a given set of solvent conditions, this denaturation (or melting) will occur at a characteristic temperature for any given DNA duplex, in proportion to the nucleotide content of the molecule, which

determines the number of hydrogen bonds stabilizing the duplex, as shown in Figure 4. If the denatured DNA is allowed to cool slowly, the duplex will again form between the two complementary strands. This process is termed renaturation or *annealing*, and is illustrated in Figure 5.

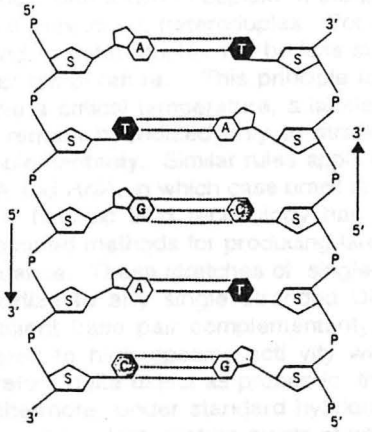


Figure 3. Antiparallel strands of the DNA double helix with hydrogen bonding between complementary bases.

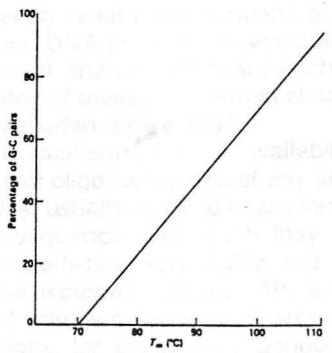


Figure 4. Relationship between G-C content and T*m* of double-stranded DNA.

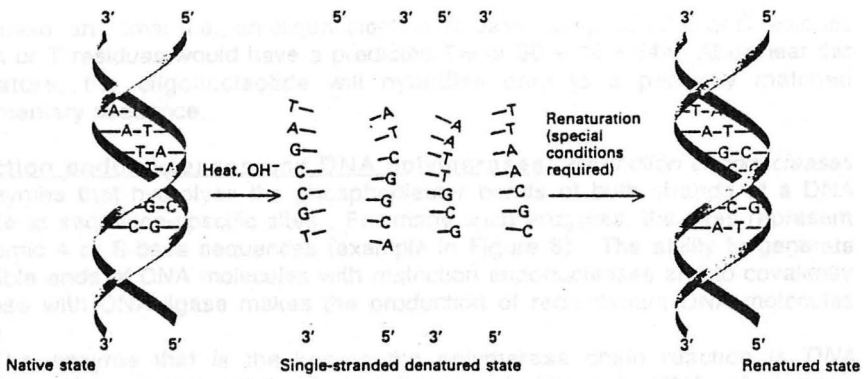


Figure 5. Denaturation and renaturation of the DNA double helix.

The same renaturation process can occur even if the two complementary strands of DNA were not originally paired or of the same length, and even if the base pairing is not exact. In this case, the process is termed *hybridization*, since the two strands form a hybrid duplex. If the base-pairing is exact, it is termed a homoduplex, and if inexact, a heteroduplex. For any given strand of DNA hybridized to another strand, a heteroduplex will be less stable than a homoduplex and will denature at a lower temperature. This principle is used to identify unknown specimens of DNA; above a critical temperature, a labeled single strand DNA probe of known specificity will remain hybridized only to strands with which it shares significant nucleotide complementarity. Similar rules apply to the formation of duplexes between strands of DNA and RNA, in which case uracil in RNA pairs with adenosine in DNA.

Nucleic acid technology has been greatly facilitated by the availability of automated methods for producing large quantitites of oligonucleotides of any desired sequence. These stretches of single-stranded DNA, usually 15 to 30 bases long, will hybridize to any single stranded DNA or RNA sequence with which they share sufficient base pair complementarity. They are chemically very stable and easily labeled to high specific acti vity with radioactive isotopes, usually ^{32}P, and are therefore quite useful as probes for the analysis of unknown nucleic acid sequences. Furthermore, under standard hybridization conditions, for any given oligonucleotide sequence, a temperature exists at which hybridization will be maintained only with a perfectly matched complementary sequence. Thus, oligonucleotide hybridization can be used to distinguish between DNA sequences that differ by as little as a single base. The melting temperature, T_m, for any oligonucleotide can be estimated by the formula:

$$T_m \text{ (centigrade)} = 4^{\circ}(G + C) + 2^{\circ}(A + T).$$

For example, an 18mer (i.e., an oligonucleotide 18 bases long) with 9 G or C residues and 9 A or T residues would have a predicted T_m of $36 + 18 = 54^{\circ}$. At or near this temperature, the oligonucleotide will hybridize only to a perfectly matched complementary sequence.

Restriction endonucleases and DNA polymerases. *Restriction endonucleases* are enzymes that hydrolyze the phosphodiester bonds of both strands of a DNA molecule at sequence-specific sites. For many such enzymes, the sites represent palindromic 4 or 6 base sequences (example in Figure 8). The ability to generate compatible ends of DNA molecules with restriction endonucleases and to covalently join these with DNA ligase makes the production of recombinant DNA molecules feasible.

The enzyme that is the key to the polymerase chain reaction is *DNA polymerase*, the principal activity of which is shown in Figure 6. DNA polymerases add nucleotides to the 3' end of a DNA strand that is growing in a 5' to 3' direction, provided that the strand is paired with an antiparallel strand that serves as a template for the addition of each new base. They are unable to initiate strands *de novo* and are therefore absolutely dependent on the presence of a primer. A number of techniques

used in molecular biology, including DNA sequencing, involve the hybridization of oligonucleotide primers to single-stranded DNA, followed by the extension of the single strand by DNA polymerase in the presence of the four nucleotide triphosphates.

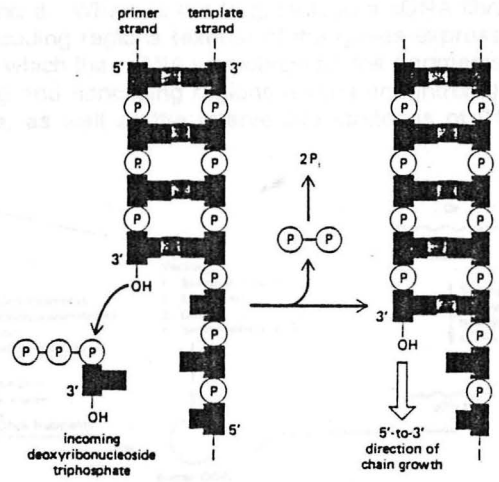


Figure 6. DNA polymerases catalyze the stepwise addition of a deoxyribonucleotide (dNTP) to the 3'-OH end of a DNA strand that is paired to an antiparallel template strand. The new strand grows in a 5'-to-3' direction. Specificity is determined by pairing of the incoming base with the template strand. The reaction is driven by a large favorable free energy change.

The most commonly used DNA polymerase is DNA polymerase I of *E. coli*. This single polypeptide has three separate activities; in addition to its 5'-->3' polymerase activity, it can degrade double-stranded DNA from a free 5' end (5'-->3' exonuclease activity) and a free 3'-OH end (3'-->5' exonuclease, or *proofreading*, activity). A proteolytic cleavage product of DNA polymerase I, called the *Klenow fragment*, lacks the 5'-->3' exonuclease activity but retains the other two activities and is the form of the enzyme most often used as a DNA polymerase in molecular biological manipulations.

Another DNA polymerase that is extremely useful in molecular biology is RNA-dependent DNA polymerase, or *reverse transcriptase*, derived from oncogenic retroviruses. This enzyme is essential in the generation of complementary DNA (cDNA) from messenger RNA (mRNA). As with other DNA polymerases, reverse transcriptase can only extend a strand that has already hybridized to a template, and thus annealing of oligonucleotide primers to the mRNA template is necessary in order to initiate cDNA first strand synthesis with reverse transcriptase.

<u>Cloning vectors and the construction of DNA libraries</u>. In order to study an individual gene at the DNA level, it is necessary to obtain reasonable quantitities of it in pure form. The human genome contains approximately 3 x 10^9 base pairs, whereas individual genes are usually contained on segments of DNA of only a few thousand base pairs. Thus, a gene that is encoded by only one locus in the genome (i.e., a

single copy gene) is only represented by about one millionth of the total human genomic DNA. Similarly, a typical cell transcibes thousands of genes into mRNA, and thus cDNA synthesized from isolated mRNA may represent several thousand different genes. Both genomic DNA and cDNA fragments can be cloned from *libraries* constructed as shown in Figures 7 and 8. Whereas the fragments in a cDNA library represent predominantly the protein-coding regions (exons) of the genes expressed as mRNA in the particular tissue from which the mRNA was obtained, the fragments in a genomic library contain both coding and noncoding regions (exons and introns) of all genes in the individual's genome, as well as the intervening stretches of DNA between genes.

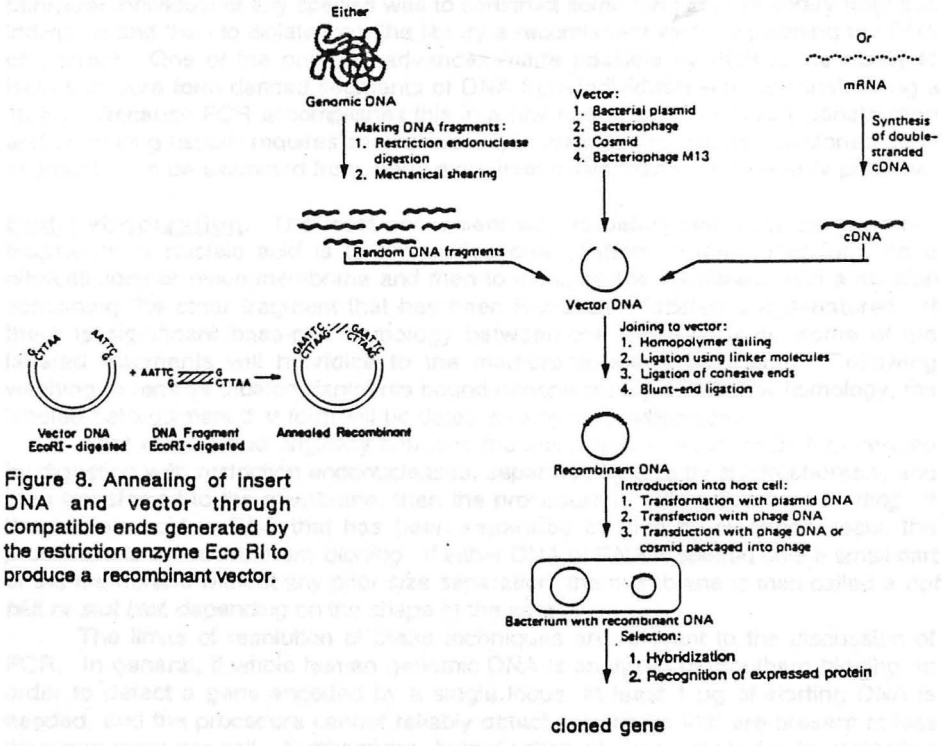


Figure 7. Classical methods of isolating cloned genes by construction of genomic and complementary DNA libraries.

Figure 8. Annealing of insert DNA and vector through compatible ends generated by the restriction enzyme Eco RI to produce a recombinant vector.

The basis of DNA libraries is the ligation of a collection of restriction fragments of DNA into a *vector*, usually a plasmid or a bacteriophage, to produce a recombinant vector that is capable of replicating when introduced into a microbial cell host. The most common hosts are certain modified strains of *E. coli*, although other hosts can be used. Usually, only a single recombinant vector molecule will transform an individual host cell, so isolation of a particular restriction fragment of DNA in a particular vector is accomplished by isolating a colony of bacteria derived from the original transformed bacterial cell. This usually involves screening thousands to tens of thousands of individual recombinant vectors by any of several methods for the DNA segment of interest.

Before the advent of PCR, the only way to clone a segment of DNA from a particular individual of any species was to construct some form of DNA library from that individual and then to isolate from the library a recombinant vector containing the DNA of interest. One of the principal advances made possible by PCR is the ability to isolate in pure form defined segments of DNA from individuals without constructing a library. Because PCR accomplishes this in a few hours, whereas library construction and screening usually requires several weeks, in practice this means that cloned gene segments can be examined from many more individuals than was previously possible.

Blot-hybridization. The most convenient way to detect homology between two fragments of nucleic acid is to immobilize one of them in denatured form on a nitrocellulose or nylon membrane and then to incubate the membrane with a solution containing the other fragment that has been isotopically labeled and denatured. If there is significant base-pair homology between the two fragments, some of the labeled fragments will hybridize to the membrane-bound fragments. Following washing to remove labeled fragments bound nonspecifically or with low homology, the labeled heterodimers that form will be detectable by autoradiography.

If the nucleic acid originally bound to the membrane consists of DNA generated by digestion with restriction endonucleases, separated by size by electrophoresis, and then transferred to the membrane, then the procedure is called *Southern blotting*. If the nucleic acid is RNA that has been separated by size by electrophoresis, the procedure is called *Northern blotting*. If either DNA or RNA is spotted onto a small part of the membrane without any prior size separation, the membrane is then called a *dot blot* or *slot blot*, depending on the shape of the spot .

The limits of resolution of these techniques are relevant to the discussion of PCR. In general, if whole human genomic DNA is analyzed by Southern blotting, in order to detect a gene encoded by a single locus, at least 1 µg of starting DNA is needed, and the procedure cannot reliably detect sequences that are present at less than one copy per cell. Furthermore, hybridization of oligonucleotides to restriction fragments generated from whole human genomic DNA is technically difficult and cannot reliably be used to detect single base differences. Hybridization of oligonucleotides to dot blots or slot blots of whole human genomic DNA are unreliable because of partial cross-hybridizations with irrelevant sequences. In contrast, hybridization of oligonucleotides to either Southern blots or dot blots of cloned DNA is

very sensitive and specific because the target fragment is present in approximately 10^6-fold greater quantity than in genomic DNA.

C. The polymerase chain reaction.

Original description with Klenow fragment of DNA polymerase. The principles of the polymerase chain reaction are shown schematically in Figure 9. In the beginning, a target DNA sequence is selected for amplification. It can be either cloned or uncloned DNA from any source, but at least part of the sequence must already be known because oligonucleotide primers complementary to the two ends of the sequence to be amplified must be available. Often, part of the sequence is not known, at least in the particular specimen to be examined.

The DNA is heat-denatured to separate the two strands. Primers complementary to the coding strand and the noncoding strand, respectively, are added and the mixture is cooled to a temperature below the lower of the two T_m for the two primers to allow the primers to hybridize to the appropriate single strands. DNA polymerase and the four deoxyribonucleotide triphosphates (dNTP) are then added and the mixture is incubated at a temperature that is optimal for the DNA polymerase in order to synthesis the complementary strand for each single stranded template. At this point, for the stretch of DNA defined by the two primers, there is twice as much double-stranded DNA as there was originally. The mixture is then heated to separate the strands, and the process of hybridization and DNA synthesis are repeated, again doubling the amount of double-stranded DNA in the region defined by the two primers. Repetetive cycles of the process eventually lead to the accumulation of a large amount of double-stranded DNA of length defined by the original two primers and of sequence identical to the starting template in that region.

In the original description of the method [1], the DNA polymerase used was the Klenow fragment of *E. coli* DNA polymerase I. Like most enzymes, this enzyme is irreversibly denatured at temperatures used to denature double-stranded DNA. Therefore, a fresh supply of this DNA polymerase needed to be added during each cycle. Furthermore, the temperature optimum for this enzyme is 37° C, which is below the T_m for many oligonucleotides. Thus, if the target DNA is uncloned genomic DNA, at this temperature the primers could hybridize to several sites on the basis of only partial homology, resulting in amplification of unwanted sequences. This problem is potentially aggravated by the 3'-->5' exonuclease activity of the Klenow fragment, which can replace mismatched bases in the primer to create a closer match to an originally only partially matched template.

Despite these limitations, including the amplification of an appreciable amount of nonspecific sequence, the original method produced specific amplifications in the range of 10^5 to 10^6-fold of sequences of several hundred base pairs [1, 3, 4] and proved useful for a number of novel applications, some of which will be discussed later. However, a significant improvement in the method was made with the introduction of a thermostable DNA polymerase, and it is this version of PCR on which most of the references to be presented are based.

The Polymerase Chain Reaction

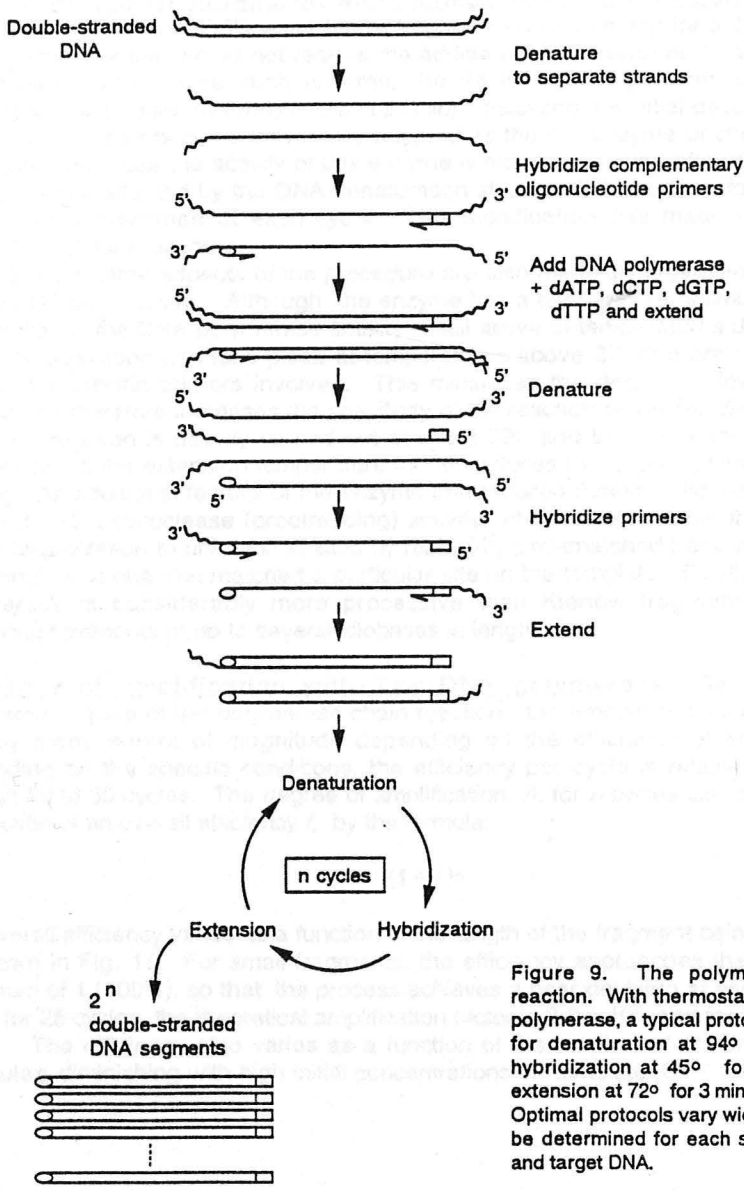


Figure 9. The polymerase chain reaction. With thermostable *Taq* DNA polymerase, a typical protocol might call for denaturation at 94° C for 1 min, hybridization at 45° for 2 min, and extension at 72° for 3 min, for 30 cycles. Optimal protocols vary widely and must be determined for each set of primers and target DNA.

The polymerase chain reaction with thermostable *Taq* DNA polymerase.

The limitations of PCR with Klenow fragment led to a search for a thermostable DNA polymerase that would not require the addition of fresh enzyme to each sample during each cycle. One such enzyme, the 94 kDA DNA polymerase from the thermophilic bacterium *Thermus aquaticus* (*Taq*), following the initial description of its use in PCR [2] has become universally adopted as the the enzyme of choice for this procedure. Because the activity of this enzyme is maintained even after incubation at 95° C, it is not affected by the DNA denaturation step of PCR and therefore does not need to be replenished at each cycle. This modification has made feasible the automation of the procedure.

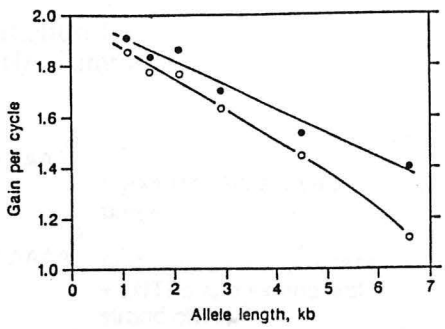
Several other aspects of the procedure are also markedly improved by the use of *Taq* DNA polymerase. Although the enzyme has a temperature optimum between 70 and 80° C, the DNA polymerase activity is still active at temperatures down to 37°. Primer hybridization can take place at temperatures above 37° that are closer to the Tm for the specific primers involved. This minimizes the degree of low-specificity priming and therefore increases the specificity of the reaction products. Since the the extension reaction is usually carried out at about 72°, and the shift from the priming temperature to the extension temperature further reduces the degree of low-specificity priming. An additional feature of the enzyme that reduces non-specific reactions is its lack of 3'-->5' exonuclease (proofreading) activity, which would enable it to promote primer hybridization to unintended sites by replacing a mismatched base at the 3' end of a primer with one that matched a particular site on the template. Finally, *Taq* DNA polymerase is considerably more processive than Klenow fragment, efficiently amplifying fragments of up to several kilobases in length.

Efficiency of amplification with *Taq* DNA polymerase. Because of the exponential nature of the polymerase chain reaction, the amount of final product can vary by many orders of magnitude depending on the efficiency of amplification. Depending on the specific conditions, the efficiency per cycle is relatively constant through 20 to 30 cycles. The degree of amplification, A, for n cycles can be estimated as function of an overall efficiency f, by the formula:

$$A = (1 + f)^n$$

The overall efficiency varies as a function of the length of the fragment being amplified, as shown in Fig. 14. For small fragments, the efficiency approaches the theoretical maximum of 1 (100%), so that the process achieves a near doubling at each cycle [5]. Thus, for 25 cycles, the theoretical amplification factor is 3.3×10^7, and for 30 cycles, 1×10^9. The efficiency also varies as a function of the number of starting template molecules, diminishing with high initial concentrations of template [6].

Figure 10. Efficiency of the polymerase chain reaction as a function of the length of the target sequence. Extension times of 6 min (open circles) or 15 min (closed circles). From reference 5.



Sensitivity and fidelity of PCR with *Taq* DNA polymerase. Initial experiments with *Taq* DNA polymerase indicated that a gene could be successfully amplified from 1 µg of genomic DNA of a cell sample in which the gene was present at a level of one target molecule in 10^6 cells [2]. Although the sensitivity varies with the particular combination of primers and target molecule, amplification of sequences present in 10^{-3} to 10^{-6} of the sampled cells have frequently been reported. Detection of PCR-amplified products at this sensitivity is approximately 10^6-fold greater than the sensitivity of conventional Southern blotting.

All DNA polymerases exhibit a finite error rate in replicating sequences. If PCR is to be useful in distinguishing small differences between DNA sequences such as point mutations, it is important that the error rate of the DNA polymerase used be acceptably low. Several estimates of the error rate of *Taq* DNA polymerase have been reported. Almost all errors involve misincorporations rather than insertion or deletions. An initial estimate of 2×10^{-4} misincorporations per nucleotide per cycle [2] now seems somewhat high. A more recent estimate shows an error rate of about 1×10^{-5} [7]. Because errors may occur with equal frequency in any of the cycles, the proportion of sequences in the final amplification product containing any particular error is usually small, and in practice only very rarely would an error be expected to confound an analysis by PCR.

Amplification of cDNA from mRNA. As will be presented in more detail in the second part of this discussion, many important questions pertinent to clinical medicine and cell biology involve identification or characterization of particular species of mRNA. Conveniently, the sensitivity and specificity of PCR have been readily adapted to the amplification of mRNA (Figure 11). The method involves isolation of RNA from the cell sample of interest, followed by the use of reverse transcriptase to synthesize a first strand cDNA. Thereafter, the procedure is similar to that used for amplifying DNA.

The Polymerase Chain Reaction for Amplifying cDNA from an mRNA Template

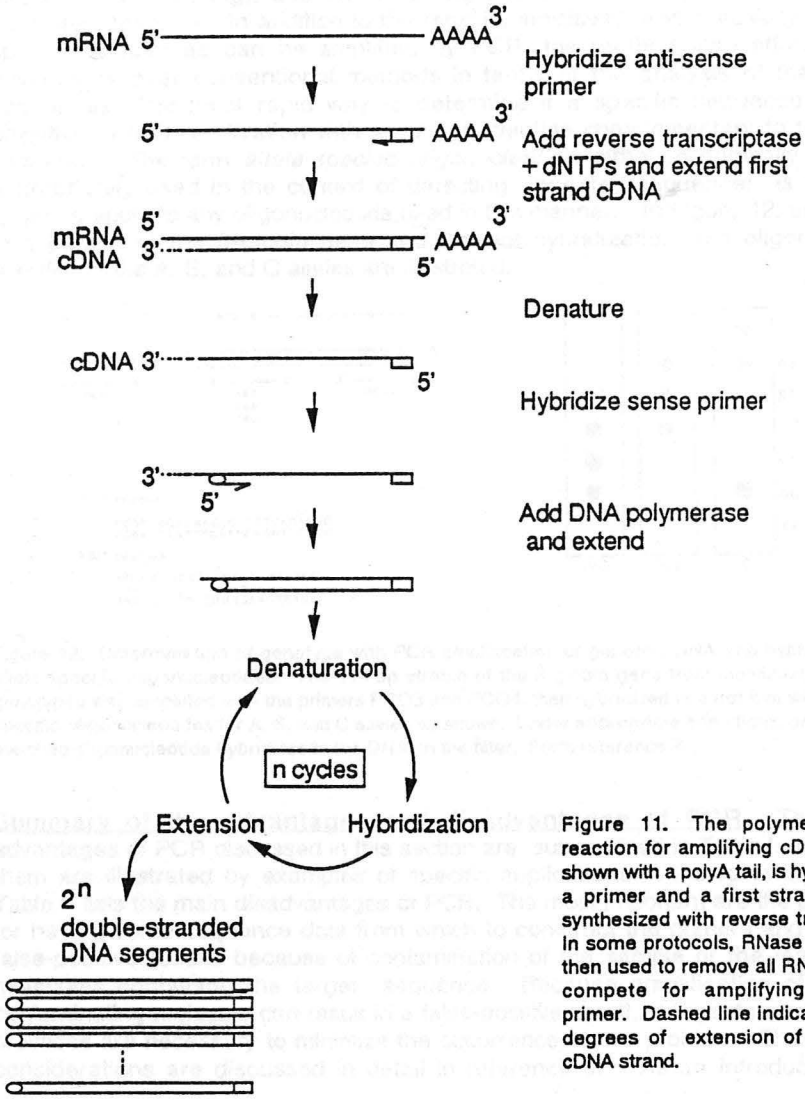


Figure 11. The polymerase chain reaction for amplifying cDNA. mRNA, shown with a polyA tail, is hybridized with a primer and a first strand cDNA is synthesized with reverse transcriptase. In some protocols, RNase treatment is then used to remove all RNA that might compete for amplifying anti-sense primer. Dashed line indicates variable degrees of extension of the original cDNA strand.

The analysis of sequences amplified by PCR.

Once segments of DNA have been cloned, they are typically analyzed in any or all of four separate ways: (a) hybridization with known probes; (b) determining sites for restriction endonucleases (restriction mapping); (c) determining the nucleotide sequence (sequencing); and (d) identifying and characterizing the transcribed and translated products. In addition to the rapidity, sensitivity, and specificity with which specific sequences can be amplified by PCR, the method also affords several advantages over conventional methods in terms of the analysis of the amplified sequences. The most rapid way to determine if a specific sequence has been amplified is by hybridization with an oligonucleotide complementary to the desired sequence. The term *allele-specific oligonucleotide*, although originally and most appropriately used in the context of detecting inherited sequences, is sometimes meant to apply to any oligonucleotide used in this manner. In Figure 12, amplification of a portion of the ß-globin gene and dot-blot hybridization with oligonucleotides specific for the A, S, and C alleles are illustrated.

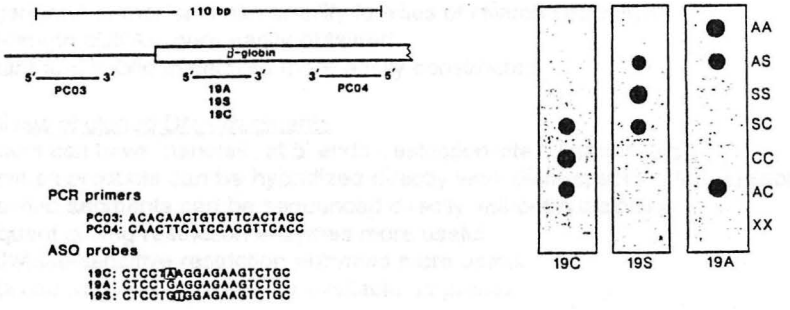


Figure 12. Determination of genotype with PCR amplification of genomic DNA and hybridization with allele specific oligonucleotides. The 110 bp stretch of the ß-globin gene from individuals of different genotypes was amplified with the primers PCO3 and PCO4, then hybridized in a dot blot with the allele-specific oligonucleotides for A, S, and C alleles as shown. Under appropriate conditions, only a perfectly matched oligonucleotide hybridizes to the DNA on the filter. From reference 3.

Summary of the advantages and disadvantages of PCR. The principal advantages of PCR discussed in this section are summarized in Table I, and many of them are illustrated by examples of specific applications in the following sections. Table II lists the main disadvantages of PCR. The most important are the prerequisite for having some sequence data from which to construct the primers and the risk of false-positive results because of contamination of the sample or the reagents with molecules containing the target sequence. Because amplification of even one contaminating molecule can result in a false-positive result, stringent precautions and practices are necessary to minimize the occurrence of this problem. These practical considerations are discussed in detail in reference 8. As an introduction to the

following section, Table III lists some general applications of PCR to clinical medicine.

Table I. Some Advantages of PCR over Conventional Methods ofRecombinant DNA Technology

Generation of cloned DNA fragments
Much less starting DNA or RNA needed
Less pure starting DNA or RNA needed
Source of DNA can be fixed or highly degraded tissue
Library not needed
RNA may come from convenient tissue
Primers can have built-in allelic or other specificity
Degenerate primer sets can amplify families of related sequences
Full-length cDNAs more easily obtained
Mutant and hybrid molecules more easily constructed
Analysis of cloned DNA fragments
Primers can have "handles" at 5' ends (restriction sites, ligands, etc.)
Amplified products can be hybridized directly with allele-specific oligonucleotides
Amplified segments can be sequenced directly without subcloning
Frequent cutting restriction enzymes more useful
Methylase-sensitive restriction enzymes more useful
Amplified fragments are rapidly available as probes

Table II. Some Disadvantages of PCR, Compared with Conventional Methods of Recombinant DNA Technology

Some sequence data from target gene must be known
Sequence must be amplifiable
Contamination is a much greater problem because of sensitivity
Obtainable fragment size is comparatively limited
Quantitation may be difficult

Table III. Prinicipal Advantages of PCR in Clincal Medicine and Clinical Investigation

Detection of DNA or RNA species present at levels of 1 copy per 10^5 to 10^6 cells
foreign sequences
somatically mutated sequences
germline sequences
Analysis of specimens at the DNA sequence level from large numbers of subjects
Analysis of DNA from fixed or partially degraded tissue: retrospective analysis
Analysis of sequences from single cells

II. APPLICATIONS OF PCR TO CLINICAL MEDICINE

A. DETECTION OF FOREIGN GENES IN HUMAN TISSUES BY PCR

A number of studies have documented that PCR-based methods can be used to detect inefectious agents in human tissues with unprecedented sensitivity, specificity, and rapidity. In individual patients, PCR has been shown to detect the presence of a number of agents, including the hepatitis B and human immunodeficiency viruses, with a sensitivity surpassing all other in vitro methods. In addition, the capacity of the procedure to be used for screening large numbers of specimens, including decades-old fixed and embedded tissue, will facilitate a variety of epidemiologic studies of the potential involvement of various agents with human disease. The following are a few examples of the application of PCR to detecting foreign genomes in human tissues.

1. Detection of the Human Immunodeficiency Virus

Detection of HIV infection is an issue of major concern. Negative serologic tests do not exclude the presence of the virus; conventional virus isolation is cumbersome, takes up to 3 to 4 weeks, and is unsuitable for routine screening. Because of the low frequency of infected cells, conventional Southern blotting is too insensitive to detect proviral sequences in samples of peripheral blood lymphocyte DNA or RNA. The sensitivity of the PCR assay has proven to be quite useful for the rapid and sensitive detection of HIV. PCR has been used to amplify DNA sequences of the HIV-1 *gag*, *pol* , *env*, and *LTR* genes, as well as RNA from the spliced transactivator *tat* gene and the 3' open reading frame [9-17]. PCR amplification of nucleic acid sequences from HIV-1-infected samples is at least six orders of magnitude more sensitive than standard nucleic acid detection methods, and several assay systems have been

described for potential clinical use [18-21]. Although still not used for routine clinical diagnosis, PCR has helped provide answers to several epidemiologic questions regarding HIV infection.

PCR vs. culture and serology. In one of the first reports of the application of PCR to detecting HIV-1 [15], proviral HIV-1 sequences were detected in 100% of DNA specimens from seropositive, homosexual men from whom the virus was isolated by coculture, and in 64% of DNA specimens from seropositive, virus culture-negative homosexual men, but in none of the DNA specimens from a control group of seronegative, virus culture-negative persons. Results were obtained within 3 days, whereas conventional virus isolation takes up to 3 to 4 weeks. Primer pairs from multiple regions of the HIV-1 genome were used to achieve maximum sensitivity.

Subsequent studies showed that PCR was not only more sensitive than virus culture, but also more sensitive than serology for detecting evidence of HIV-1 infection. In one report [22], four asymptomatic homosexual men were studied who reverted from seropositive to seronegative for HIV-1 over 2.5 years, as shown by both ELISA and Western blot, and whose cryopreserved peripheral blood mononuclear cells were negative for HIV-1 by standard culture assay. HIV-1 provirus was detected by PCR in all four subjects 6 to 18 months after the last positive antibody test. Thus, asymptomatic persons seropositive for HIV-1 may not remain seropositive, but may remain latently infected with HIV-1, as detectedby PCR.

PCR has more recently been used in examining the latent period following HIV infection before serum antibodies are detectable [11]. A cohort was studied of 133 seronegative homosexual men who continued to be involved in high-risk sexual activity. HIV-1 was cultured from PBM from 31 of the 133 men, 27 of whom remained seronegative for up to 36 months after the positive culture. The other four men seroconverted 11 to 17 months after the isolation of HIV-1. In three of them, PCR analysis of cryopreserved lymphocytes demonstrated that HIV-1 provirus had been present 23, 35, and 35 months before seroconversion.

Detection of HIV-1 sequences in brain tissue by PCR has also been reported in a case of acute fatal CNS infection in a patient lacking serologic evidence of HIV infection [23].

HIV infection in infants and children. Unlike the problem in adults of a latent period before seroconversion, serology in newborn children of HIV-seropositive mothers is of no value in detecting infection because of the transmission and persistence of maternal antibodies. In a large cooperative study [24], PCR was used to detect HIV infection in infants of seropositive women. during the neonatal (age less than 28 days) and postneonatal periods. HIV-1 sequences were found in five of seven neonates and six of six postneonatal infants in whom AIDS later developed, but not in nine neonates or 25 infants who remained well, or in 15 infants born to HIV-seronegative mothers. The data from this study are shown in Table IV. The authors concluded that PCR will apparently be useful in diagnosing HIV infection in newborns and in predicting the subsequent development of AIDS. In another study

[10], HIV-1 sequences were detected by PCR in blood samples from six of 14 newborn infants of seropositive mothers, as well as in five of 10 seronegative older children born of HIV-1-infected mothers.

Table IV. From reference 24.

Detection of Proviral Sequences of HIV by Polymerase Chain Reaction in Infants Born to HIV-Seropositive Mothers, According to the Clinical Status of the Infants and Their Age at the Time the Blood Sample Was Obtained.

Most Recent Clinical Status of Infant	Time of Testing	
	Neonatal Period	Postneonatal Period
	infants positive/infants tested	
III with AIDS* (n = 8)	5/7	6/6
III with nonspecific findings suggestive of HIV infection (n = 14)	1/8	4/14
Clinically well		
Followed ≥15 mo (n = 18)	0/6	0/18
Followed <15 mo (n = 47)†	0/3	0/7

*As defined by the CDC.[11]

†Most available samples from this group will be tested when follow-up samples have been obtained.

HIV in household or sexual contacts. A few studies have applied PCR to the question of detecting HIV infection in seronegative individuals exposed to known HIV-infected contacts. In one study [16], PCR was used to confirm the presence of HIV viral DNA detected by in situ hybridization in leukocytes from seronegative children of HIV-infected mothers and in three seronegative sexual partners of seropositive drug addicts. In contrast to these findings, a much lower rate of HIV infection was found in 23 healthy women with histories of repeated unprotected sexual exposure to HIV-1 infected hemophiliacs [12], who were studied for HIV-1 antibody, serum antigen, HIV-1 DNA sequences by PCR, and HIV-1 virus isolation from peripheral mononuclear cells. Twenty-two of 23 (96%) women were negative by all HIV-1 assays. One woman was positive by all the HIV-1 assays, including ELISA. These preliminary results, employing the sensitivity of PCR, suggest that the frequency of HIV-1 infection in antibody-negative sexual partners of HIV-1 infected hemophiliacs is probably very low. The same authors had previously demonstrated by direct viral culture that HIV-1 seropositivity in this group of hemophiliacs represented active infection, not simply an immune response to non-infectious material contained in cryoprecipitates [25]. Thus, studies at the level of PCR sensitivity have confirmed earlier suggestions that the rate of transmission of HIV-1 to sexual contacts of seropositive drug abusers is much higher than the rate of transmission of HIV-1 to sexual contacts of seropositive hemophiliacs.

HIV and other coexistent retroviral infections. Because of serologic cross-reactivity among different retroviruses, in some cases serology may not be reliable in detecting true infection. Alternatively, the coexistence of another infection may alter the clinical presentation of HIV infection. HIV-2, which has been described

most frequently in Western Africa, is serologically cross-reactive with HIV-1. In a seroprevalence study of 944 persons in Abidjan, Cote d'Ivoire, West Africa [17], one person dually reactive for HIV-1 and HIV-2 by various serological techniques was found to harbor proviral sequences from both viruses, as detected by PCR. The authors noted that this was is the first confirmed case of a mixed HIV-1 and HIV-2 infection in a single individual.

In another report [26], a patient was described with a histologically benign, polyclonal HTLV-I infection exhibiting both an absolute CD4+ and CD8+ lymphocytosis. HTLV-I+ pulmonary tract nodules remitted completely with alkylating agents and steroids, but subsequently, the patient developed a severe immunodeficiency state and expired. Retrospective serologic and PCR assays for HIV-1 demonstrated that he had been doubly infected from the time of presentation. Postmortem analysis by PCR revealed the presence of both HTLV-I and HIV-1 in lymphatic tissues and the testes, and HIV-1 was also detected in brain tissue.

Table V summarizes some of the many reports of the clinical application of PCR in detecting HIV-1 sequences.

Table V. Detection of HIV-1 in Clinical Specimens by PCR

Source of Specimens	References
Patients	
Patients negative by virus isolation	[15]
Patients negative by serology	[11, 22, 23]
Infants of HIV-1-infected mothers	[10, 24]
Coexistent HIV-2 infection	[17]
Coexistent HTLV-1 infection	[26]
Objects in contact with patients	
Endoscopes	[27]
Leukocyte filters	[28]

2. Detection of hepatitis B virus by PCR

Several studies have documented detection of hepatitis B virus by PCR in serum of infected individuals at a sensitivity exceeding all previous methods, including DNA hybridization [29-31]. Ulrich et al. [32], comparing the level of detection of PCR with the level of transmissibility in serum samples that had been standardized for infectivity of chimpanzees, found that PCR was about10-fold more sensitive than

transmissibility for detecting HBV. A corollary of this finding is that samples truly negative by PCR are probably not infectious. Thiers, et al. [33] studies three patients negative for all HBV serological markers. HBV sequences were amplified by PCR from serum in all three patients. Serum from two of the patients was injected into chimpanzees that subsequently developed hepatitis. HBV sequences were then identified in PCR-amplified samples from the chimpanzees. Given these results, it will be of interest to examine serum and liver biopsy material from patients with chronic liver disease in the absence of serologic markers for HBV.

3. Specific detection of Herpesvirus family members by PCR.

Several studies have documented the rapid, sensitive and specific detection of members of the Herpesvirus family. Cytomegalovirus (CMV) has been detected by PCR in blood, fresh or paraffin-fixed tissue, and newborn urine [34-37]. In all of these studies, the assay for CMV by PCR appeared to have significant clinical utility. One study using PCR found no correlation between the presence of CMV genomes and the appearance of Kaposi's sarcoma in patients with AIDS [38].

Other studies have reported the detection of HHV-6 (human herpes virus 6, human B-lymphotropic virus) has been identified by PCR in peripheral blood, tumor tissue, and retina [39, 40], Herpes simplex virus in paraffin-embedded biopsy specimens [41], and increased levels of Epstein-Barr virus in salivary gland tissue and peripheral blood from patients with Sjögren's syndrome [42].

4. HTLV-1 and neurologic disease

Human T-cell lymphotropic virus type I (HTLV-1) is a retrovirus that was originally isolated from cases of T-cell lymphoma and there is strong evidence that it is causally associated with adult human T cell leukemia/lymphoma [43]. Recently, a high proportion of patients with tropical spastic paraparesis, a chronic neurologic disorder characterized by low back pain, bladder dysfunction, and a progressive spastic paraparesis, have been found to have antibodies that react with HTLV-1. Several reports have now documented that a retrovirus is definitely associated with tropical spastic papaparesis, and either direct isolation of proviral DNA [44] or PCR amplification and sequencing [45-47] have identified the retrovirus as HTLV-1.

Two recent studies have used PCR to address the question of a retroviral association with multiple sclerosis. Greenberg, et al. [48] examined DNA from peripheral blood mononuclear cells from 21 patients with chronic progressive multiple sclerosis. Six patients showed amplification of sequences homologous to the HTLV-1 *pol* gene, and three of these also showed amplification of *env* sequences, but none showed amplification of any other HTLV-1 sequence. None of 35 controls showed any positive reactions. Reddy et al. [49] detected both *gag* and *env* amplified HTLV-1sequences in six of six patients with MS. These were confirmed by DNA sequencing. The intensity of amplification was considerably greater when adherent peripheral blood mononuclear cells were used as the DNA source, rather than unseparated cells. One of 20 normal control samples was positive in this study. A

potential weakness of both of these last two studies is the absence of a control group with other neurologic diseases. Nonetheless, although these results are preliminary, they suggest that retroviral infection may play a role in the pathogenesis of MS, and they point the way toward further studies involving large numbers of patients.

5. Distinguishing HTLV-II from HTLV-I. Confirmed infection with HTLV-II has been difficult to establish because of substantial serologic cross-reactivity between HTLV-II and HTLV-I. Lee et al. [50] used PCR amplification of proviral sequences to study i.v. drug abusers in New Orleans who were seropositive for HTLV-I Of 23 individuals confirmed by PCR analysis to be infected with either HTLV, 21 were found to be infected with HTLV-II, and only two with HTLV-I. With PCR identification of the virus, it will now be possible to to investigate the pathogenicity of HTLV-II in humans.

6. Human papillomavirus in cervical neoplasia. Certain types of human papillomavirushave been linked epidemiologically to squamous cell carcinoma of the uterine cervix and the cervical dysplasia that precedes it. Two specific types of HPV, 16 and 18, have been most closely associated with cervical carcinoma, but these have also been found to be widely distributed in the normal population, so the significance of the association with malignancy has been unclear. A number of studies examining either fresh or fixed and embedded tissue have used PCR amplification to detect type-specific HPV genomes [51-57]. Although type 16 was commonly found in dysplastic or malignant tissue, it was also found in a proportion of normal specimens. Of particular interest is a recent report [57] that used PCR to amplify a 313 bp sequence of the upstream regulatory region HPV 16. In some cases, the amplified sequence gave a band that appeared somewhat smaller than predicted when visualized by agarose gel electrophoresis. The shorter band was found by DNA sequencing to contain a 21 base deletion, and the subtype from which it was derived was termed 16b, whereas the standard sequence subtype was termed 16a. Only the presence of HPV 16a correlated strongly with cervical carcinoma and intraepthelial neoplasia (Table VI). Although in need of confirmation and further investigation, this finding is potentially very significant in defining the role of oncogenic HPV in cervical neoplasia.

7. Detection of Mycobacterium by PCR

The diagnosis of mycobacterial infections remains a frequent clinical problem. Direct identification of acid-fast organisms is rapid but insensitive both quantitatively and qualitatively. Culture methods take many weeks and may require multiple specimens to be positive. A recent report suggests that PCR may be useful in the detection and identification of mycobacterial organisms in clinical specimens. Mycobacteria share a common 65 kD antigen that shows considerable shared homology among the different species, with some regions of variability. Hance et al. [58] successfully amplified a 383 bp sequence from the DNA encoding this protein in four different mycobacterial species, including *M. tuberculosis* and *M. avium*. These sequences were then distinguished by hybridization with species-specific

oligonucleotide probes. Although no clinical specimens were analyzed, it was demonstrated that as few as 3 mycobacterial organisms mixed with one million human peripheral blood mononuclear cells could be detected after DNA extraction, amplification, and hybridization.

Table VI. Human Papillomavirus Types 16a and 16b in Uterine Cervical Neoplasia

Subjects	HPV16a only	HPV16b only	both
Normal (n = 141)	1	89	28
Normal post laser therapy (n = 20)	1	17	2
Dyskaryotic (n = 30)	13	6	2
Carcinoma (n = 22)	22	0	0

From reference [57].

A summary of the infectious agents discussed in this section is listed in Table VII.

Table VII. Detection of Foreign Genes in Human Tissues by PCR

Retroviruses	Other viruses	Bacteria
HIV-1	Hepatitis B	Mycobacteria
HIV-2	Herpesviruses	
HTLV-I	CMV	
HTLV-II	HSV-1	
	HHV-6	
	EBV	
	Papillomavirus	

B. DETECTION OF SOMATIC MUTATIONS BY PCR

Detection of somatically acquired mutations presents specially problems because they usually occur in only a small fraction of the cells of a particular tissue. The speed and sensitivity of PCR in detecting a mutated sequence present in very low abundance, as well as the capacity of the technique to amplify DNA from fixed, embedded specimens, has proven to be particularly advantageous in studying both normal and malignant cells for genetic alterations. The most prominent example of physiologic somatic gene alteration is the receptor gene rearrangement that gives rise to the diversity of immunoglobulin and T-cell receptor molecules in the immune response, and several studies have documented the usefulness of PCR in cloning these rearranged genes from small numbers of cells [59-61].

With regard to human disease, somatic mutation is most prominently linked to neoplasia, and particularly to the activation of protooncogenes. Activation of protooncogenes has been demonstrated to occur by several mechanisms, including endogenous amplification of the gene, point mutations and chromosomal rearrangements. In vitro amplification by PCR of a suspected target sequence from a small tumor sample has permitted the detection of the latter two types of mutations. PCR is particularly useful in the detection of somatic mutations in large numbers of samples. In identifying somatic mutations, the sensitivity of PCR in detecting a particular DNA or RNA species that occurs in only a small fraction of the cells being tested is also critical.

In this discussion, studies of the diseases associated with the Philadelphia chromosome will be presented as an example of the use of PCR in detecting and characterizing chromosomal rearrangements, while studies of diseases associated with the activation of *ras* protooncogenes will be used as examples of the detection of somatically acquired point mutations by PCR.

1. Chromosomal translocations.

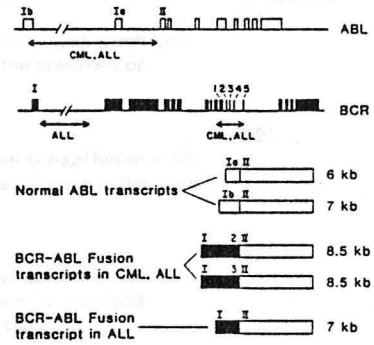
Many human neoplasms are known to be associated with chromosomal translocations, some of which have been shown to result in the activation of protooncogenes. Although detection of these rearrangements can be made by cytogenetic and Southern blot analysis, nucleotide sequence analysis of the fusion products of the rearranged genes has generally required the construction and screening of libraries from individual malignancies. Other methods of detecting mRNA species arising from hybrid genes are technically difficult and generally too cumbersome for application to large numbers of samples. Several groups have recently applied PCR techniques to the characterization of the mRNA of fusion genes resulting from chromosomal translocations.

A large number of studies have documented the sensitivity of PCR amplification of chromosomal translocation-induced hybrid genes in identifying residual leukemia or lymphoma cells (minimal resiudual disease), even in patients who have been in

studies have examined the breakpoint regions themselves. Most attention in this regard has been directed toward the products of the translocation that produces the Philadelphia (Ph') chromosome.

The Ph' chromosome results from a reciprocal translocation involving the long arms of chromosomes 9 and 22, t(9;22)(q34;q11) and has been found in over 90% of patients with chronic myelogenous leukemia (CML) and 10-25% of patients with acute lymphoblastic leukemia (reviewed in [69]). As a result of the translocation, a 3' portion of the *abl* protooncogene, normally located on chromosome 9, becomes attached to the 5' portion of the *bcr* (breakpoint cluster region) gene on chromosome 22. The chimeric *bcr-abl* gene is transcribed to produce a fusion mRNA that is translated into a protein thought to play a role in the pathogenesis of the resulting leukemias. The genes and resulting mRNA species are demonstrated schematically in Figure 13. In about half of the rearrangements in ALL, exon II of the *abl* gene is joined to the major cluster region of the *bcr* gene, shown as exons 1-5 in Fig. 13. In the other cases of ALL, the *bcr* breakpoint occurs within a 50 kb stretch 3' of exon I.

Figure13 from reference 69



Schematic representation of the normal *BCR* and *ABL* genes and the *ABL*-related transcripts present in normal and leukemic cells. (*Upper*) The *BCR* and *ABL* genes (not to scale). Exons of the *ABL* gene are depicted as open boxes above the line; the alternative first exons Ib and Ia and the *ABL* second exon are indicated by Roman numerals. Exons of the *BCR* gene are depicted as solid boxes above the line. Arabic numbers denote the exons of the bcr; the Roman numeral I denotes the first exon of the *BCR* gene. (*Lower*) *ABL*-related transcripts present in cells containing the Ph^1 chromosome. The mRNAs are depicted schematically, and their approximate sizes are given in kb. For each transcript, the coding sequences derived from the *BCR* gene are shown as solid boxes, using the numbering system as in *Upper*.

By selecting primers corresponding to the appropriate 5' portions of the *bcr* gene and 3' portions of the *abl* gene, several groups of investigators have amplified cDNAs from the *bcr-abl* fusion transcripts of leukemia cells from patients with CML or ALL. The results of these studies are summarized in Table VIII. The studies to date have confirmed previous findings that several different *bcr-abl* mRNA species occur in CML and in ALL, including cases in which the the Ph' chromosome is cytogenetically absent. There is no evidence for a previously suspected second translocation involving *bcr*-abl fusion in the blast crisis of CML. As in other systems, PCR appears to be the most sensitive way to detect residual leukemic cells following treatment, both in bone marrow and in peripheral blood. Sequencing of PCR-amplified cDNA from a large number *bcr-abl* fusion genes will undoubtedly be accomplished in the near future and should contribute significantly to better understanding of the intracellular

function of the resulting fusion proteins, as well as to improved clinical classification of the *bcr-abl* fusion-positive leukemias.

In contrast to the heterogeneity found in Ph'-positive leukemias, another study employing PCR found a striking homogeneity among follicular lymphomas associated with the t(14;18) translocation [70]. In this case, the site of interest was the minor cluster region (mcr) flanking the bcl-2-proto-oncogene, previously shown to be a frequent site of breakpoints. Eight of 10 mcr+ breakpoints clustered within a 500 nucleotide region, and five occurred within three nucleotides of each other. Because this translocation is found so consistently, it can detected by direct amplification of genomic DNA with primers that flank the breakpoint region, without the need for mRNA amplification.

Table VIII. Studies of Ph' *bcr-abl* Fusion Genes in CML and ALL by PCR

Principal Findings	Ref.
Target: mRNA from CML and ALL Results: Diagnosis of the *bcr-abl* translocation by PCR is rapid, much more sensitive than other protocols, and independent of the presence or absence of an identifiable Ph chromosome	[71]
Target: mRNA from 8 Ph+ *bcr*- ALL Results: Confirms that in all 8 cases studied, the novel *bcr-abl* fusion in Ph+ *bcr*- ALL is the result of joining of the first exon of the *bcr* gene to the c-*abl* oncogene.	[72]
Target: mRNA from CML in blast crisis Results: No evidence for alternative splicing of the *bcr-abl* mRNA in two CML blast crisis lines; does not support the hypothesis for a second mutational event involving the *bcr-abl* region in blast crisis	[73]
Target: mRNA from 20 patients with classic CML and 1 with ALL Results: *bcr* exon 3 present in14 leukemias and absent in 7; documents the heterogeneity of the fusion product	[74]
Target: mRNA from 4 patients with Ph' negative CML Results: *bcr-abl* translocation productive of mRNA can result from complex rearrangements involving any of several different chromosomes	[75]
Target: mRNA from CML patient blood and bone marrow before and after myeloablative therapyResults: *bcr-abl* mRNA is still expressed even after myeloablative therapy in preparation for bone marrow transplantation	[76]

Target: *bcr-abl* mRNA from 11 cases of Ph'-positive ALL, including 2 representing a lymphoblast crisis progression of CML. [69]
Results: 1 of 6 *bcr+* cases and 1 of 5 *bcr-* cases contained none of the known fusion mRNA species, while another 1 case contained both bcr+ and bcr- fusion species. These findings demonstrate unexpected heterogeneity within the subtypes of Ph' ALL.

Target: mRNA from blood samples from 31 patients with Ph'-positive CML and two patients with Ph'-negative *bcr* rearranged CML. [66]
Findings: Additional evidence for heterogeneity of hybrid mRNAs, both in chronic phase and in blast crisis. All eight patients in complete cytogenetic remission had evidence of minimal residual Ph'-positive clones detected by PCR.

---

2. Point mutations in *ras* genes.

Introduction [77, 78]. The family of *ras* genes has been a focus of intense research interest since 1982, when these genes were first found to be associated with human tumors. The name is derived from the term *rat sarcoma* because these genes were first identified as the transforming principle of two strains of rat sarcoma viruses. *ras* genes are highly conserved throughout eukaryotic evolution. In all species examined, they encode proteins that bind guanine nucleotides, have GTPase activity, are associated with cell plasma membranes, and have homology with G proteins. These and other properties suggest that they are intimately involved with transmembrane signalling and the control of cellular proliferation. Three functional *ras* genes are found in the human genome, designated H-*ras*-1, K-*ras*-2, and N-*ras*, encoded, respectively, on the short arms of chromosomes 11, 12, and 1, and referred to hereafter as H-*ras*, K-*ras*, and N-*ras*. Two alternative forms of the K-*ras*-encoded protein result from the use of two alternative fourth exons. The three functional human *ras* genes encode closely related proteins generically termed p21, each with four domains. The homology of these protein sequences to each other and to *ras* genes of rat, mouse, chicken, fruit fly, slime mold, and yeast is shown in Figure 14.

Although the precise intracellular function of normal *ras* proteins has yet to be elucidated, two striking properties with respect to human neoplasia have been found. First, *ras* genes acquire in vitro and in vivo transforming properties as a result of specific, somatically acquired point mutations; that is, the normal *ras* genes are proto-oncogenes, whereas the transforming mutated genes are oncogenes. Second, *ras* oncogenes have been found in a wide variety of human malignancies, and have been estimated to have an overall incidence of 10-15% of all human cancers. Mutations in naturally occurring *ras* oncogenes have been localized to amino acid positions 12, 13, 59, and 61. Of these, mutations at 12, 13, and 61 have been associated with human malignancies. Substitutions at these positions are associated with varying degrees of oncogenesis. Thus, substitution of virtually any other amino

acid for the normal glycine residue at position 12 or the normal glutamine at position 61 is associated with oncogeneic activation of the *ras* proteins.

Although the identification and characterization of *ras* oncogenes in human tumors have been actively pursued, much of the original work relied primarily on relatively insensitive gene transfer experiments or technically difficult nucleic acid hybridization experiments. Not surprisingly, until recently no strong correlations had been made between any particular *ras* oncogene and any particular human tumor. The sensitivity and selectivity of PCR for specific sequences, as well as the ability to analyze fixed, paraffin-embedded specimens, have facilitated a number of studies of different tumor types, from which some coherent findings have begun to emerge. These will be considered on a tumor-by-tumor basis.

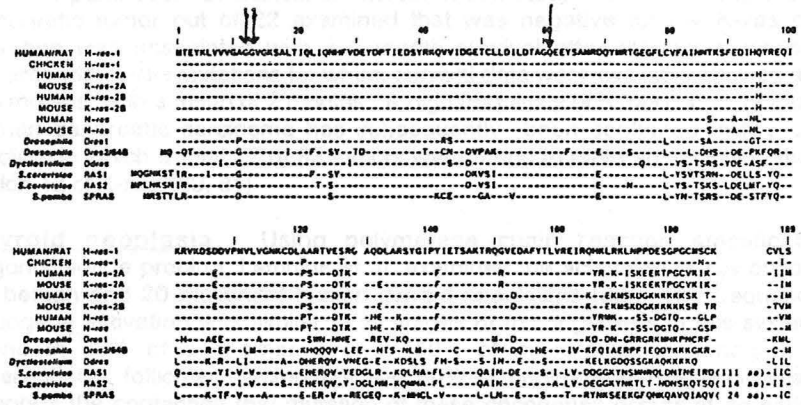


Figure 14. The protein sequences of *ras* genes from several species, shown in the one letter amino acid code. Activating mutations occur in the absolutely conserved glycine (G) residues at positions 12 and 13 and glutamine (Q) residue at position 61, indicated with vertical arrows. From reference 77.

Human malignancies showing frequent association with *ras* mutations by PCR analysis.

Carcinoma of the exocrine pancreas. Almoguera et al. [79], using PCR in conjunction with the RNase A mismatch cleavage method to detect mutations, examined K-*ras* genes in human carcinomas of the exocrine pancreas. Examining both frozen tumor specimens and single 5μm sections from formalin-fixed, paraffin-embedded tumor tissue surgically removed or obtained at autopsy, they found that 21 out of 22 pancreatic carcinomas contained K-*ras* genes with mutations at codon 12. In all seven cases tested, the mutation was present in both primary tumors and their corresponding metastases. No mutations were detected in normal tissue, including normal pancreatic tissue from the same cancer patients, or in five gallbladder carcinomas. From these results, it appears that K-*ras* somatic mutational activation is a critical event in the oncogenesis of most, if not all, human cancers of the exocrine pancreas. Of additional interest in this study was the finding that the one pancreatic tumor out of 22 examined that was negative for the K-*ras* codon 12 mutation was associated with a 9 month survival after diagnosis, while10 other patients with K-*ras* mutations for whom survival data were available showed a range of 1-4 months, with a mean of 2 months. A high frequency of K-*ras* codon 12 mutations in human pancreatic carcinoma has subsequently been confirmed in two additional studies, in which a total 77 of 93 tumors were found to have an activating mutation of codon 12 of K-*ras* [80, 81].

Thyroid neoplasia. Using polymerase chain reaction amplification and oligonucleotide probing, Lemoine et al. examined the activation of *ras* oncogenes in 24 benign and 20 malignant human thyroid neoplasms [82]. The frequency of *ras* oncogene activation was similar at all stages of tumorigenesis in this system, being found in 33% of adenomas overall (50% of microfollicular tumours), 53% of differentiated follicular carcinomas and 60% of undifferentiated carcinomas. This supports the contention that mutation of these oncogenes occurs at an early step in tumorigenesis. The predominant amino acid substitution in the differentiated tumours was glutamine to arginine at position 61 of H-*ras* or N-*ras*, but this mutation was not found in any of the undifferentiated tumours. It was noted that while transition mutations predominated in differentiated tumours (both benign and malignant), transversions were more common in the undifferentiated tumours.

Colon carcinoma. Burmer et al. [83, 84] studied paraffin-embedded colon tumors for activating mutations of K-*ras*. Twenty-six of 40 carcimomas showed mutations in codon 12. In one sample in which both carcinoma and adjacent villous adenoma were analyzed, the same K-*ras* mutation was detected in both tissues, but not in the adjacent normal tissue. Nuclei were obtained from paraffin-embedded colon carcinomas, sorted into aneuploid and diploid subpopulations by flow cytometry, and then analyzed for mutations K-*ras* gene by direct sequencing of PCR amplified product. Mutations in codon 12 of K-*ras* were present in both aneuploid and

diploidsubpopulations of sorted carcinomas. These findings suggest that mutations in K-ras precede ploidy alterations and may be an early event in the progression of colon carcinoma.

Hematologic malignancies. A number of studies have been conducted looking for point mutations of *ras* genes in hematologic malignancies, primarily leukemias and myelodysplastic syndromes. Mutations have most commonly been found in N-*ras* in acute myelogenous leukemia (AML) and in myeoldysplasia with or without leukemic transformation. Table IX contains a summary of the most informative studies reported.

Malignancies not associated with ras mutations. Several malignancies have been found at the level of PCR detection not to be associated with activating *ras* mutations, including esophageal squamous cell carcinoma [91], hepatocelluar carcioma [92], breast cancer [93], CLL and non-Hodgkin's lymphoma [85].

Table IX. *ras* Mutations in Hematologic Malignancies

Malignancies studied	Findings	Reference
Lymphoid malignancies (178): B- and T-cell ALL, CLL, non-Hodgkin lymphoma studied for activating mutations involving codons 12 and 61 of H-*ras*, K-*ras* and N-*ras*, and codon 13 of the N-*ras*.	Mutations found in codons 12 or 13 of N-*ras* in 6 of 33 cases of acute lymphoblastic leukemia (18%), correlating with the most undifferentiated phenotype. No mutations found in CLL or NHL. Five of the 6 ALL cases had the same single base substitution.	[85]
Acute myeloid leukemia (AML): 52 presentation AML DNAs studied for mutations in codons 12, 13, and 61 of N-*ras* and in codons 12 and 61 of K-*ras* and H-*ras*.	Mutations found in 14 (27%) -all in N-*ras* and predominantly in codon 12. No correlation between disease subtype and the incidence or type of N-*ras* mutation. In four patients, 2 N-*ras* mutations were found to coexist, arising in different N-*ras* alleles. A mutated N-*ras* gene was absent in four patients in relapse in whom it had been detected at presentation.	[86]
Primary myelodysplasia: 34 cases studied for mutations in codons 12, 13, and 61 of H-*ras*, K-*ras*, and N-*ras*.	Mutations at codon 12 of K-*ras* or N-*ras* were detected in three cases (9%), all detectable in peripheral blood. No correlation with progression to leukemic phase.	[87]

Primary myelodysplasia: 50 cases studied for mutations in codons 12, 3,61 of H-, K-, and N-*ras*.	Mutations in codons 12/13 and 61 of *ras* genes were detected in 20 patients (40%). Two patients had mutations detected in two different *ras* genes. The prevalence of mutations was highest in the subgroup with CMML.	[88]
Myelodysplasia (19), AML (10), CML (10), ALL (2), CLL (2) studied for mutations in N-*ras*.	Mutations in N-*ras* found in 6 cases: 3 MDS, 3 AML. One positive case with AML was in complete remission.	[89]
Three cases of idiopathic myelofibrosis studied for the presence of mutations *ras* codons 12, 13, or 61.	One patient showed a point mutation at codon 12 of N-*ras* in peripheral blood. The N-*ras* mutation was found in granulocytes, monocytes, B and T lymphocytes, as well as erythroblasts. This suggests that idiopathic myelofibrosis originated from a pluripotent stem cell in this patient.	[90]

Table X. Summary of Frequent *ras* Mutations Identified by PCR in Human Malignancies

Condition	Mutated *ras* gene	% of cases
Carcinoma of exocrine pancreas	K-*ras*	75 - 95
Adenocarcinoma of the colon	K-*ras*	65
Thyroid carcinoma	H-, N-, K-*ras*	50 - 60
Acute myelogenous leukemia	N-*ras*	27 - 30
Acute lymphoblastic leukemia	N-*ras*	18
Myelodysplasia	H-, N-, K-*ras*	10 - 40

Table XI. Summary of Human Malignancies Found by PCR to be Infrequently Associated with *ras* Mutations

Condition	Frequency of *ras* Mutations
Ovarian carcinoma	0/37
Esophageal squamous cell carcinoma	0/41
Hepatocellular carcinoma	2/39
Breast carcinoma	1/40
Chronic lymphocytic leukemia	0/51
Non-Hodgkin's lymphoma	0/88

References

1. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA , Arnheim N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350-4 (1985).

2. Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB , Erlich HA. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-91 (1988).

3. Saiki RK, Bugawan TL, Horn GT, Mullis KB , Erlich HA. Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. Nature 324: 163-6 (1986).

4. Embury SH, Scharf SJ, Saiki RK, Gholson MA, Golbus M, Arnheim N , Erlich HA. Rapid prenatal diagnosis of sickle cell anemia by a new method of DNA analysis. N Engl J Med 316: 656-61 (1987).

5. Jeffreys AJ, Wilson V, Neumann R , Keyte J. Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. Nucleic Acids Res 16: 10953-71 (1988).

6. Syvanen AC, Bengtstrom M, Tenhunen J , Soderlund H. Quantification of polymerase chain reaction products by affinity-based hybrid collection. Nucleic Acids Res 16: 11327-38 (1988).

7. Williams JF. Optimization strategies for the polymerase chain reaction.

BioTechniques 7: 762-769 (1989).

8. Kwok S , Higuchi R. Avoiding false positives with PCR. Nature 339: 237-8 (1989).

9. Hart C, Schochetman G, Spira T, Lifson A, Moore J, Galphin J, Sninsky J , Ou CY. Direct detection of HIV RNA expression in seropositive subjects. Lancet 2: 596-9 (1988).

10. Laure F, Courgnaud V, Rouzioux C, Blanche S, Veber F, Burgard M, Jacomet C, Griscelli C , Brechot C. Detection of HIV1 DNA in infants and children by means of the polymerase chain reaction. Lancet 2: 538-41 (1988).

11. Imagawa DT, Lee MH, Wolinsky SM, Sano K, Morales F, Kwok S, Sninsky JJ, Nishanian PG, Giorgi J, Fahey JL , et al. Human immunodeficiency virus type 1 infection in homosexual men who remain seronegative for prolonged periods. N Engl J Med 320: 1458-62 (1989).

12. Jackson JB, Kwok SY, Hopsicker JS, Sannerud KJ, Sninsky JJ, Edson JR , Balfour HHJ. Absence of HIV-1 infection in antibody-negative sexual partners of HIV-1 infected hemophiliacs. Transfusion 29: 265-7 (1989).

13. Kumar R, Goedert JJ , Hughes SH. A method for the rapid screening of human blood samples for the presence of HIV-1 sequences: the probe-shift assay. Aids Res Hum Retroviruses 5: 345-54 (1989).

14. Murakawa GJ, Zaia JA, Spallone PA, Stephens DA, Kaplan BE, Wallace RB , Rossi JJ. Direct detection of HIV-1 RNA from AIDS and ARC patient samples. Dna 7: 287-95 (1988).

15. Ou CY, Kwok S, Mitchell SW, Mack DH, Sninsky JJ, Krebs JW, Feorino P, Warfield D , Schochetman G. DNA amplification for direct detection of HIV-1 in DNA of peripheral blood mononuclear cells [published erratum appears in Science 1988 Apr 8;240(4849):240]. Science 239: 295-7 (1988).

16. Pezzella M, Rossi P, Lombardi V, Gemelli V, Mariani CR, Mirolo M, Fundaro C, Moschese V , Wigzell H. HIV viral sequences in seronegative people at risk detected by in situ hybridisation and polymerase chain reaction. Bmj 298: 713-6 (1989).

17. Rayfield M, De CK, Heyward W, Goldstein L, Krebs J, Kwok S, Lee S, McCormick J, Moreau JM, Odehouri K , et al. Mixed human immunodeficiency virus (HIV) infection in an individual: demonstration of both HIV type 1 and type 2 proviral sequences by using polymerase chain reaction. J Infect Dis 158: 1170-6 (1988).

18. Keller GH, Huang DP , Manak MM. A sensitive nonisotopic hybridization assay for HIV-1 DNA. Anal Biochem 177: 27-32 (1989).

19. Kwoh DY, Davis GR, Whitfield KM, Chappelle HL, DiMichele LJ , Gingeras TR. Transcription-based amplification system and detection of amplified human immunodeficiency virus type 1 with a bead-based sandwich hybridization format. Proc Natl Acad Sci U S A 86: 1173-7 (1989).

20. Guatelli JC, Gingeras TR , Richman DD. Nucleic acid amplification in vitro: detection of sequences with low copy numbers and application to diagnosis of human immunodeficiency virus type 1 infection. Clin Microbiol Rev 2: 217-26 (1989).

21. Kemp DJ, Smith DB, Foote SJ, Samaras N , Peterson MG. Colorimetric detection of specific DNA segments amplified by polymerase chain reactions. Proc Natl Acad Sci U S A 86: 2423-7 (1989).

22. Farzadegan H, Polis MA, Wolinsky SM, Rinaldo CRJ, Sninsky JJ, Kwok S, Griffith RL, Kaslow RA, Phair JP, Polk BF , et al. Loss of human immunodeficiency virus type 1 (HIV-1) antibodies with evidence of viral infection in asymptomatic homosexual men. A report from the Multicenter AIDS Cohort Study. Ann Intern Med 108: 785-90 (1988).

23. Varma VA, Hunter S, Tickman R, Srinivasan A , Swan D. Acute fatal HIV encephalitis with negative serologic assays for antibody and antigen: diagnosis by polymerase chain reaction [letter]. N Engl J Med 320: 1494-5 (1989).

24. Rogers MF, Ou CY, Rayfield M, Thomas PA, Schoenbaum EE, Abrams E, Krasinski K, Selwyn PA, Moore J, Kaul A , et al. Use of the polymerase chain reaction for early detection of the proviral sequences of human immunodeficiency virus in infants born to seropositive mothers. New York City Collaborative Study of Maternal HIV Transmission and Montefiore Medical Center HIV Perinatal Transmission Study Group. N Engl J Med 320: 1649-54 (1989).

25. Jackson JB, Sannerud KJ, Hopsicker JS, Kwok SY, Edson JR , Balfour HHJ. Hemophiliacs with HIV antibody are actively infected. Jama 260: 2236-9 (1988).

26. Ehrlich GD, Davey FR, Kirshner JJ, Sninsky JJ, Kwok S, Slamon DJ, Kalish R , Poiesz BJ. A polyclonal CD4+ and CD8+ lymphocytosis in a patient doubly infected with HTLV-I and HIV-1: a clinical and molecular analysis. Am J Hematol 30: 128-39 (1989).

27. Hanson PJ, Gor D, Clarke JR, Chadwick MV, Nicholson G, Shah N, Gazzard B, Jeffries DJ, Gaya H , Collins JV. Contamination of endoscopes used in AIDS patients. Lancet 2: 86-8 (1989).

28. Rawal BD, Busch MP, Endow R, Garcia dLJ, Perkins HA, Schwadron R , Vyas GN. Reduction of human immunodeficiency virus-infected cells from donor blood by leukocyte filtration. Transfusion 29: 460-2 (1989).

29. Larzul D, Guigue F, Sninsky JJ, Mack DH, Brechot C , Guesdon JL. Detection of hepatitis B virus sequences in serum by using in vitro enzymatic amplification. J Virol Methods 20: 227-37 (1988).

30. Sumazaki R, Motz M, Wolf H, Heinig J, Jilg W , Deinhardt F. Detection of hepatitis B virus in serum using amplification of viral DNA by means of the polymerase chain reaction. J Med Virol 27: 304-8 (1989).

31. Kaneko S, Miller RH, Feinstone SM, Unoura M, Kobayashi K, Hattori N , Purcell RH. Detection of serum hepatitis B virus DNA in patients with chronic hepatitis using the polymerase chain reaction assay. Proc Natl Acad Sci U S A 86: 312-6 (1989).

32. Ulrich PP, Bhat RA, Seto B, Mack D, Sninsky J , Vyas GN. Enzymatic amplification of hepatitis B virus DNA in serum compared with infectivity testing in chimpanzees. J Infect Dis 160: 37-43 (1989).

33. Thiers V, Nakajima E, Kremsdorf D, Mack D, Schellekens H, Driss F, Goudeau A, Wands J, Sninsky J, Tiollais P , et al. Transmission of hepatitis B from hepatitis-B-seronegative subjects. Lancet 2: 1273-6 (1988).

34. Demmler GJ, Buffone GJ, Schimbor CM , May RA. Detection of cytomegalovirus in urine from newborns by using polymerase chain reaction DNA amplification. J Infect Dis 158: 1177-84 (1988).

35. Shibata D, Martin WJ, Appleman MD, Causey DM, Leedom JM , Arnheim N. Detection of cytomegalovirus DNA in peripheral blood of patients infected with human immunodeficiency virus. J Infect Dis 158: 1185-92 (1988).

36. Cassol SA, Poon MC, Pal R, Naylor MJ, Culver JJ, Bowen TJ, Russell JA, Krawetz SA, Pon RT , Hoar DI. Primer-mediated enzymatic amplification of cytomegalovirus (CMV) DNA. Application to the early diagnosis of CMV infection in marrow transplant recipients. J Clin Invest 83: 1109-15 (1989).

37. Chehab FF, Xiao X, Kan YW , Yen TS. Detection of cytomegalovirus infection in paraffin-embedded tissue specimens with the polymerase chain reaction. Mod Pathol 2: 75-8 (1989).

38. Van den Berg F, Schipper M, Jiwa M, Rook R, Van de Rijke F , Tigges B. Implausibility of an aetiological association between cytomegalovirus and Kaposi's

sarcoma shown by four techniques. J Clin Pathol 42: 128-31 (1989).

39. Buchbinder A, Josephs SF, Ablashi D, Salahuddin SZ, Klotman ME, Manak M, Krueger GR, Wong SF , Gallo RC. Polymerase chain reaction amplification and in situ hybridization for the detection of human B-lymphotropic virus. J Virol Methods 21: 191-7 (1988).

40. Qavi HB, Green MT, SeGall GK , Font RL. Demonstration of HIV-1 and HHV-6 in AIDS-associated retinitis. Curr Eye Res 8: 379-87 (1989).

41. Cao M, Xiao X, Egbert B, Darragh TM , Yen TS. Rapid detection of cutaneous herpes simplex virus infection with the polymerase chain reaction. J Invest Dermatol 92: 391-2 (1989).

42. Saito I, Servenius B, Compton T , Fox RI. Detection of Epstein-Barr virus DNA by polymerase chain reaction in blood and tissue biopsies from patients with Sjogren's syndrome. J Exp Med 169: 2191-8 (1989).

43. Yoshida M , Seiki M. Recent advances in the molecular biology of HTLV-1: trans-activation of viral and cellular genes. Ann. Rev. Immunol. 5: 541-59 (1987).

44. Reddy EP, Mettus RV, DeFreitas E, Wroblewska Z, Cisco M , Koprowski H. Molecular cloning of human T-cell lymphotrophic virus type I-like proviral genome from the peripheral lymphocyte DNA of a patient with chronic neurologic disorders. Proc Natl Acad Sci U S A 85: 3599-603 (1988).

45. Bangham CR, Daenke S, Phillips RE, Cruickshank JK , Bell JI. Enzymatic amplification of exogenous and endogenous retroviral sequences from DNA of patients with tropical spastic paraparesis. Embo J 7: 4179-84 (1988).

46. Desgranges C, Bechet JM, Couderc LJ, Caubarrere I , Vernant JC. Detection of HTLV-1 DNA by polymerase chain reaction in alveolar lymphocytes of patients with tropical spastic paraparesis [letter]. J Infect Dis 160: 162-3 (1989).

47. Saxton EH, Lee H, Swanson P, Chen IS, Ruland C, Chin E, Aboulafia D, Delamarter R , Rosenblatt JD. Detection of human T-cell leukemia/lymphoma virus type I in a transfusion recipient with chronic myelopathy. Neurology 39: 841-4 (1989).

48. Greenberg SJ, Ehrlich GD, Abbott MA, Hurwitz BJ, Waldmann TA , Poiesz BJ. Detection of sequences homologous to human retroviral DNA in multiple sclerosis by gene amplification. Proc Natl Acad Sci U S A 86: 2878-82 (1989).

49. Reddy EP, Sandberg WM, Mettus RV, Ray PE, DeFreitas E , Koprowski H.

Amplification and molecular cloning of HTLV-I sequences from DNA of multiple sclerosis patients. Science 243: 529-33 (1989).

50. Lee H, Swanson P, Shorty VS, Zack JA, Rosenblatt JD , Chen IS. High rate of HTLV-II infection in seropositive i.v. drug abusers in New Orleans. Science 244: 471-5 (1989).

51. Shibata DK, Arnheim N , Martin WJ. Detection of human papilloma virus in paraffin-embedded tissue using the polymerase chain reaction. J Exp Med 167: 225-30 (1988).

52. Shibata D, Fu YS, Gupta JW, Shah KV, Arnheim N , Martin WJ. Detection of human papillomavirus in normal and dysplastic tissue by the polymerase chain reaction. Lab Invest 59: 555-9 (1988).

53. Morris BJ, Flanagan JL, McKinnon KJ , Nightingale BN. Papillomavirus screening of cervical lavages by polymerase chain reaction [letter]. Lancet 2: 1368 (1988).

54. Dallas PB, Flanagan JL, Nightingale BN , Morris BJ. Polymerase chain reaction for fast, nonradioactive detection of high- and low-risk papillomavirus types in routine cervical specimens and in biopsies. J Med Virol 27: 105-11 (1989).

55. Young LS, Bevan IS, Johnson MA, Blomfield PI, Bromidge T, Maitland NJ , Woodman CB. The polymerase chain reaction: a new epidemiological tool for investigating cervical human papillomavirus infection. Bmj 298: 14-8 (1989).

56. McDonnell JM, Mayr AJ , Martin WJ. DNA of human papillomavirus type 16 in dysplastic and malignant lesions of the conjunctiva and cornea. N Engl J Med 320: 1442-6 (1989).

57. Tidy JA, Vousden KH , Farrell PJ. Relation between infection with a subtype of HPV16 and cervical neoplasia. Lancet 1: 1225-7 (1989).

58. Hance AJ, Grandchamp B, Levy-Frebault V, Lecossier D, Rauzier J, Bocart D , Gicquel B. Detection and identification of mycobacteria by amplification of mycobacterial DNA. Molec. Microbiol. 3: 843-849 (1989).

59. Lacy MJ, McNeil LK, Roth ME , Kranz DM. T-cell receptor delta-chain diversity in peripheral lymphocytes. Proc Natl Acad Sci U S A 86: 1023-6 (1989).

60. Larrick JW, Danielsson L, Brenner CA, Abrahamson M, Fry KE , Borrebaeck CA. Rapid cloning of rearranged immunoglobulin genes from human hybridoma cells using mixed primers and the polymerase chain reaction. Biochem Biophys Res

Commun 160: 1250-6 (1989).

61. Orlandi R, Gussow DH, Jones PT , Winter G. Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. Proc Natl Acad Sci U S A 86: 3833-7 (1989).

62. Crescenzi M, Seto M, Herzig GP, Weiss PD, Griffith RC , Korsmeyer SJ. Thermostable DNA polymerase chain amplification of t(14;18) chromosome breakpoints and detection of minimal residual disease. Proc Natl Acad Sci U S A 85: 4869-73 (1988).

63. dAuriol L, Macintyre E, Galibert F , Sigaux F. In vitro amplification of T cell gamma gene rearrangements: a new tool for the assessment of minimal residual disease in acute lymphoblastic leukemias. Leukemia 3: 155-8 (1989).

64. Janssen JW, Buschle M, Layton M, Drexler HG, Lyons J, van dBH, Heimpel H, Kubanek B, Kleihauer E, Mufti GJ , et al. Clonal analysis of myelodysplastic syndromes: evidence of multipotent stem cell origin. Blood 73: 248-54 (1989).

65. Lee MS, Chang KS, Cabanillas F, Freireich EJ, Trujillo JM , Stass SA. Detection of minimal residual cells carrying the t(14;18) by DNA sequence amplification. Science 237: 175-8 (1987).

66. Lee MS, Chang KS, Freireich EJ, Kantarjian HM, Talpaz M, Trujillo JM , Stass SA. Detection of minimal residual bcr/abl transcripts by a modified polymerase chain reaction. Blood 72: 893-7 (1988).

67. Lee MS, LeMaistre A, Kantarjian HM, Talpaz M, Freireich EJ, Trujillo JM , Stass SA. Detection of two alternative bcr/abl mRNA junctions and minimal residual disease in Philadelphia chromosome positive chronic myelogenous leukemia by polymerase chain reaction. Blood 73: 2165-70 (1989).

68. Yamada M, Hudson S, Tournay O, Bittenbender S, Shane SS, Lange B, Tsujimoto Y, Caton AJ , Rovera G. Detection of minimal disease in hematopoietic malignancies of the B-cell lineage by using third-complementarity-determining region (CDR-III)-specific probes. Proc Natl Acad Sci U S A 86: 5123-7 (1989).

69. Hooberman AL, Carrino JJ, Leibowitz D, Rowley JD, Le BMM, Arlin ZA , Westbrook CA. Unexpected heterogeneity of BCR-ABL fusion mRNA detected by polymerase chain reaction in Philadelphia chromosome-positive acute lymphoblastic leukemia. Proc Natl Acad Sci U S A 86: 4259-63 (1989).

70. Ngan BY, Nourse J , Cleary ML. Detection of chromosomal translocation t(14;18) within the minor cluster region of bcl-2 by polymerase chain reaction and

direct genomic sequencing of the enzymatically amplified DNA in follicular lymphomas. Blood 73: 1759-62 (1989).

71. Kawasaki ES, Clark SS, Coyne MY, Smith SD, Champlin R, Witte ON , McCormick FP. Diagnosis of chronic myeloid and acute lymphocytic leukemias by detection of leukemia-specific mRNA sequences amplified in vitro. Proc Natl Acad Sci U S A 85: 5698-702 (1988).

72. Hermans A, Gow J, Selleri L, von Lindern M, Hagemeijer A, Wiedemann LM , Grosveld G. bcr-abl oncogene activation in Philadelphia chromosome-positive acute lymphoblastic leukemia. Leukemia 2: 628-33 (1988).

73. Hermans A, Selleri L, Gow J , Grosveld GC. Absence of alternative splicing in bcr-abl mRNA in chronic myeloid leukemia cell lines. Blood 72: 2066-9 (1988).

74. Dobrovic A, Trainor KJ , Morley AA. Detection of the molecular abnormality in chronic myeloid leukemia by use of the polymerase chain reaction. Blood 72: 2063-5 (1988).

75. van der Plas DC, Hermans AB, Soekarman D, Smit EM, de Klein A, Smadja N, Alimena G, Goudsmit R, Grosveld G , Hagemeijer A. Cytogenetic and molecular analysis in Philadelphia negative CML. Blood 73: 1038-44 (1989).

76. Lange W, Snyder DS, Castro R, Rossi JJ , Blume KG. Detection by enzymatic amplification of bcr-abl mRNA in peripheral blood and bone marrow cells of patients with chronic myelogenous leukemia. Blood 73: 1735-41 (1989).

77. Barbacid M. ras Genes. Ann. Rev. Biochem. 56: 779-827 (1987).

78. Bos JL. The ras gene family and human carcinogenesis. Mutat. Res. 195: 255-271 (1988).

79. Almoguera C, Shibata D, Forrester K, Martin J, Arnheim N , Perucho M. Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. Cell 53: 549-54 (1988).

80. Grunewald K, Lyons J, Frohlich A, Feichtinger H, Weger RA, Schwab G, Janssen JW , Bartram CR. High frequency of Ki-ras codon 12 mutations in pancreatic adenocarcinomas. Int J Cancer 43: 1037-41 (1989).

81. Smit VT, Boot AJ, Smits AM, Fleuren GJ, Cornelisse CJ , Bos JL. KRAS codon 12 mutations occur very frequently in pancreatic adenocarcinomas. Nucleic Acids Res 16: 7773-82 (1988).

82. Lemoine NR, Mayall ES, Wyllie FS, Williams ED, Goyns M, Stringer B , Wynford TD. High frequency of ras oncogene activation in all stages of human thyroid tumorigenesis. Oncogene 4: 159-64 (1989).

83. Burmer GC, Rabinovitch PS , Loeb LA. Analysis of c-Ki-ras mutations in human colon carcinoma by cell sorting, polymerase chain reaction, and DNA sequencing. Cancer Res 49: 2141-6 (1989).

84. Burmer GC , Loeb LA. Mutations in the KRAS2 oncogene during progressive stages of human colon carcinoma. Proc Natl Acad Sci U S A 86: 2403-7 (1989).

85. Neri A, Knowles DM, Greco A, McCormick F , Dalla FR. Analysis of RAS oncogene mutations in human lymphoid malignancies. Proc Natl Acad Sci U S A 85: 9268-72 (1988).

86. Farr CJ, Saiki RK, Erlich HA, McCormick F , Marshall CJ. Analysis of RAS gene mutations in acute myeloid leukemia by polymerase chain reaction and oligonucleotide probes. Proc Natl Acad Sci U S A 85: 1629-33 (1988).

87. Lyons J, Janssen JW, Bartram C, Layton M , Mufti GJ. Mutation of Ki-ras and N-ras oncogenes in myelodysplastic syndromes. Blood 71: 1707-12 (1988).

88. Padua RA, Carter G, Hughes D, Gow J, Farr C, Oscier D, McCormick F , Jacobs A. RAS mutations in myelodysplasia detected by amplification, oligonucleotide hybridization, and transformation. Leukemia 2: 503-10 (1988).

89. Mano H, Ishikawa F, Hirai H , Takaku F. Mutations of N-ras oncogene in myelodysplastic syndromes and leukemias detected by polymerase chain reaction. Jpn J Cancer Res 80: 102-6 (1989).

90. Buschle M, Janssen JW, Drexler H, Lyons J, Anger B , Bartram CR. Evidence for pluripotent stem cell origin of idiopathic myelofibrosis: clonal analysis of a case characterized by a N-ras gene mutation. Leukemia 2: 658-60 (1988).

91. Hollstein MC, Smits AM, Galiana C, Yamasaki H, Bos JL, Mandard A, Partensky C , Montesano R. Amplification of epidermal growth factor receptor gene but no evidence of ras mutations in primary human esophageal cancers. Cancer Res 48: 5119-23 (1988).

92. Tsuda H, Hirohashi S, Shimosato Y, Ino Y, Yoshida T , Terada M. Low incidence of point mutation of c-Ki-ras and N-ras oncogenes in human hepatocellular carcinoma. Jpn J Cancer Res 80: 196-9 (1989).

93. Rochlitz CF, Scott GK, Dodson JM, Liu E, Dollbaum C, Smith HS , Benz CC. Incidence of activating ras oncogene mutations associated with primary and metastatic human breast cancer. Cancer Res 49: 357-60 (1989).

C. USE OF PCR IN STUDIES OF ALLELES ASSOCIATED WITH INHERITED SINGLE GENE DISORDERS.

The first application of PCR was the detection of the beta-hemoglobin allele associated with sickle cell disease [1-3], and subsequently, numerous human alleles have been studied by PCR at either the cDNA or genomic DNA levels. In discussing the application of PCR to the detection of human alleles, examples will be presented illustrating the power of this method to extend knowledge gained from conventional molecular biology.

1. Results from classical molecular genetics.

In recent years, a great many single gene disorders have been characterized at the DNA level. This has provided a wealth of new information about the structure of the defective protein products, as well as significant insights into the function of the normal protein products encoded at these loci. In general, initial cloning of a defective gene has involved either previous isolation of a cDNA encoding a normal version of the gene (normal allele) or direct isolation of the defective cDNA (disease-associated allele). In either case, to obtain the disease allele, it has been necessary to construct and screen a cDNA library derived from the appropriate tissue in an affected pedigree. Characterization of the gene defect at the level of genomic DNA has then usually required constructing and screening a genomic DNA library from the affected pedigree.

Once a cDNA probe is available for a particular genetic locus, it is often possible to deduce the genotype of an individual if the relevant alleles at that locus can be distinguished by restriction fragment length polymorphisms (RFLP) identified by Southern blotting of genomic DNA cut with the appropriate restriction enzyme and probed with the appropriate cDNA. However this procedure is too insensitive for many genes, especially those cases in which the alleles differ only by nucleotide substitutions and not by deletions of more than a few base pairs. This insensitivity arises because many nucleotide substitutions do not result in altered restriction sites recognized by enzymes that generate fragments of sizes that are distinguishable by Southern blotting.

In some cases in which RFLP derived from the disease gene itself are not available, RFLP may be found at loci that are chromosomally located near the disease locus and are therefore genetically linked. If the genotype of one member of a family is known, the genotype of other family members can often be inferred by the pattern of segregation of RFLP at loci linked to the locus in question, taking into account the probability of a meiotic recombination event between loci. In the cases of some disease loci such as those encoding the disease genes for cystic fibrosis and Huntington's disease, the actual locus encoding the disease gene has not been identified, but the genetic status of individuals in affected families can be predicted with high probability because of the existence of closely linked polymorphic loci. In these latter cases, however, screening in populations of unrelated individuals is of little or no value in identifying carriers of the disease gene.

Many common single gene diseases have been found to be heterogeneous at the DNA level. That is, several different disease alleles may exist at the same locus, often causing defective gene products by several different mechanisms. The genes involved are usually large, with many intron sequences and often with large mRNA transcripts, and therefore with many sites for potential defect-inducing mutations. A few examples of such genes are those encoding the clotting factors VIII and IX, the low-density lipoprotein receptor, the muscle protein dystrophin, and the various procollagen molecules. Identification of the many different defective DNA sequences that may underlie the diseases affecting these genes is made difficult at the genomic level by the size and complexity of the genes, and at the cDNA level by the difficulty in obtaining full-length mRNA for cDNA libraries from the affected tissues of numerous affected pedigrees.

Given the foregoing, it is not surprising that the characterization of each new mutant allele at the level of the DNA sequence has generally been the result of a significant investment of time, effort, and expense on the part of the investigator. Furthermore, screening for mutant alleles at the DNA level has generally not been practical except in the setting of prenatal diagnosis of high risk pregnancies or in families of affected individuals.

As emphasized in the introduction, PCR provides a means for examining specific DNA sequences in individuals without first having to construct a library from that individual because, in effect, the specific primers select the gene before, rather than after, the work of cloning gets underway. Several other aspects of PCR also facilitate the analysis of large numbers of alleles; these were listed in Table I and will be illustrated by the specific examples that follow. The analysis of large numbers of disease-associated alleles by PCR serves at least four distinct goals: (a) increasing knowledge of the function of the affected gene or gene product, (b) increasing knowledge of the behavior of mutant alleles within given populations, (c) screening populations for carriers of mutant alleles, and (d) prenatal diagnosis of genetic diseases. All of these are illustrated by the examples that follow.

As will be seen, PCR has already proven useful as an extremely rapid, sensitive, and specific technique for prenatal screening and carrier detection.

2. ß-Thalassemia: a single gene disease caused by many alleles.

ß-thalassemia (ß-thal) is a heterogeneous inherited disorder of ß-globin synthesis. The disease has a high frequency in populations stemming from the Mediterranean basin, Africa, and southern Asia. Approximately 50 different mutations have been described, although many fewer mutations occur within any given population (Figure 15). Genetic screening and prenatal diagnosis are useful in preventing new cases of the disease, but have been complicated by the diversity of mutations and by the fact that most ß-thal are due to point mutations that affect mRNA transcription, processing, or translation. Because point mutations are usually not detectable by conventional Southern blotting techniques, detection of ß-thal genes has until recently involved determining linkage to ß-globin haplotypes by Southern

blotting of genomic DNA from members of affected families. Several publications have recently demonstrated the utility of PCR in the direct identification of known ß-thal genes, as well as in the identification of previously undescribed ß-thalassemia mutations.

Cai et al. [5, 6] used PCR and oligonucleotide hybridization to examine ß-thal genes in the Guangdong province in southern China, a region where multiple mutations are known to occur and the carrier rate for ß-thal is estimated to be ~3%. Segments of the ß-thal gene were amplified, and dot blot hybridization of the amplified DNA with oligonucleotide probes corresponding to the six mutations previously found in southern China directly identified the mutations causing ß-thal in affected families. Exact prenatal molecular diagnoses were made in all 20 fetuses studied over a 6-month period. Representative determinations are shown in Figure 16.

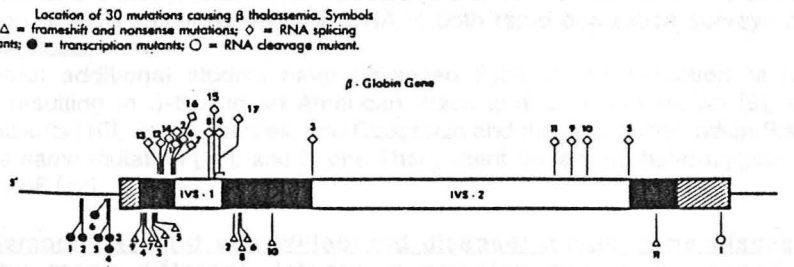


Figure 15, from reference 4.

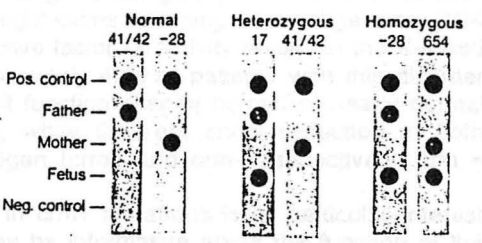


Figure 16. Prenatal screening of PCR-amplified DNA for ß-thal genes with allele specific oligonucleotides. The oligos used matched the mutant genes of each set of parents. Three cases indicate the threedifferent fetal genotypes. From reference 6.

In addition, a seventh, previously undetermined, mutation was identified in this study by direct sequencing of amplified DNA from one heterozygous parent whose amplified ß-globin DNA failed to hybridize with any of the mutant oligonucleotide probes in the original panel. In another study of patients in South China with ß-thalassemia intermedia (mild disease requiring fewer transfusions) [7], seven mutations were

identified, five of which matched those described by Cai, et al.

A similar approach was used [8] to delineate the ß-globin sequences causing ß-thal in Spain, a country that has been inhabited by different Mediterranean populations over the centuries. Fifty-eight ß-thal alleles were studied either by hybridization with allele-specific radiolabeled oligonucleotide probes or by direct sequence analysis of the amplification product. Seven different ß-thal mutations were found, all of which had been previously described in Mediterranean populations. The predominant (64%) mutation was one that has also been found in Sardinia, Morocco, Algeria, Tunisia, Greece, and Turkey, whereas the mutation most commonly causing ß-thal in the eastern part of the Mediterranean basin was underrepresented (8.5%). These findings probably reflect the presence in Spain of North African populations from the 8th through the 15th century. A first-trimester prenatal diagnosis of ß-thal was also reported in this study, which demonstrates further the usefulness of the dot-blot hybridization of PCR-amplified genomic DNA in both rapid population surveys and prenatal diagnosis.

Several additional studies have employed PCR in the detection of new mutations resulting in ß-thal in an American Black and an Asian Indian [9], two Chinese patients [10], in two families, one Caucasian and the other West Indian Black, sharing the same mutation [11], and in one Thai patient compound heterozygous for ß-thal and HbE [12].

3. The hemophilias and von Willebrand disease: single gene diseases caused by many different defects in complex genes expressed in relatively inaccessible tissues.

Hemophilia B. Human factor IX is a single-chain glycoprotein of 415 residues synthesized in the liver and encoded by eight exons spanning 34 kb of genomic DNA on the X chromosome. Absent or defective factor IX activity results in the X-linked bleeding disorder hemophilia B. Approximately 40% of patients with this disorder display decreased factor IX procoagulant function despite normal or nearly normal levels of factor IX antigen in plasma, while the rest show reduction in both procoagulant activity and factor IX antigen (crm+ and crm-, respectively; crm = *c*ross-*r*eactive *m*aterial).

Identifying the molecular defects in crm+ mutations is of particular interest because the sites of such mutations may be informative about the function of the molecule. However, there are many different mutants, the gene is large and complex, and the mRNA is relatively inaccessible; therefore, the effort involved in identifying mutant sequences in the gene by conventional means is considerable.

Several groups have recently applied PCR technology to the study of factor IX. Spitzer et al. [13] used PCR to confirm the results of conventional cloning and sequencing techniqes that identified a single base substitution at position 390 in the catalytic domain of variant factor IXBmLE. Similarly, Tsang et al. [14] used PCR to confirm the presence of a single base substitution at position 333 in the catalytic domain of a nonfunctional crm+ factor IX gene mutation (factor IX London 2).

Denton et al. [15] studied two patients with crm+ factor IX hemophilia B whose factor IX molecules failed to react with a monoclonal antibody that binds the epidermal growth factor (EGF)-like domains of the molecule. The exons that encode the first and second EGF-like domains of factor IX were amplified from the patients' genomic DNA and sequenced. Both patients had identical mutations in a highly conserved residue, probably resulting in a change in the tertiary structure of first the EGF-like domain. (A similar approach was used by Dr. H. Hobbs and her colleagues here at Southwestern to localize a deletion mutation in an exon of the low-density lipoprotein receptor gene [16]).

Carrier detection and prenatal screening are important in the prevention of the X-linked hemophilias. Conventional methods involve identification of a series of linked restriction fragment length polymorphisms, require Southern blotting of genomic DNA, and detect approximately 80% of carriers. Several groups have used PCR in combination with additional restriction enzyme sites or with other techniques to improve the sensitivity of this method in cases in which the actual molecular defect has not been identified [17-19].

Hemophilia A. Most of the arguments described above for factor IX also apply to factor VIII, deficiency of which results in hemophilia A. To date, two reports have appeared describing the use of PCR to detect defects in factor VIII [20, 21].

Platelet disorders. von Willebrand disease (vWD) is the most common inherited bleeding disorder in humans. It can result from either a quantitative or a qualitative defect of von Willebrand factor (vWF), a multimeric plasma glycoprotein that serves as a carrier for factor VIII and as a mediator of platelet adhesion to the subendothelium. vWF is synthesized in megakaryocytes and in endothelial cells. The gene encoding vWF spans at least 150 kb and contains at least 37 exons. The mRNA is very large, 8.7 kb, and post-translational processing is complex, including the asembly of a series of high-molecular mass multimers.

Molecular studies of vWD have been limited by the large size of the vWF gene and difficulty in obtaining vWF mRNA from patients. Recently, Ginsburg et al. [22] reported amplifying and sequencing vWF mRNA from peripheral blood platelets. Three separate mutations were identified: a silent vWF allele, resulting from a defect in vWF mRNA transcription or processing, and two different but adjacent missense mutations, the locations of which may identify an important vWF functional domain.

The capacity of PCR to amplify mRNA from platelets should prove generally useful for studying other platelet-specific proteins, as well as vWF. Other reports of PCR-amplified, platelet-derived mRNA have also appeared [23, 24].

4. Collagen genes: further examples of single gene diseases caused by heterogeneous defects in complex genes.

Spondyloepiphyseal dysplasias (SED) are a heterogeneous group of inherited disorders characterized by disproportionate short stature and pleiotropic involvement

of the skeletal and ocular systems. Evidence has suggested that some cases of SED may result from structural defects in type II collagen. Elucidation of the defective gene sequences in these disorders has been limited by the size and complexity of the genes involved and by the inaccessibility of chondrocytes for making cDNA libraries. Recently, Lee et al. [25] identified by Southern blotting an abnormal restriction pattern in a type II collagen gene (COL2A1) in one of the affected members of an SED kindred. The defect was found to be localized to a segment in genomic DNA containing exons 45-52, which encode the carboxy-terminal propeptide and the last 123 amino acid residues of the triple helical domain. PCR analysis of four overlapping genomic fragments within this region precisely localized the molecular defect and demonstrated that all affected family members carried the same heterozygous deletion of exon 48, confirmed by DNA sequencing. This mutation would be predicted to result in type II collagen homotrimers containing an interstitial deletion of 36 amino acids in the triple helical domain of one or more of the chains. More subtle defects in type II collagen genes may underlie a host of musculoskeletal disorders. The ability to amplify specific regions of these genes directly from genomic DNA should help accelerate progress in this area.

Defective type I collagen genes result in osteogenesis imperfecta (OI), a complex group of inherited disorders with a spectrum of disorders of bone and connective tissue. Two recent reports have documented new mutations resulting in, respectively, mild, dominantly inherited type I OI [26], and in perinatal lethal type II OI [27]. In both cases, biochemical evidence from the patients' type I collagen provided information that led to amplification of selected regions of the respective a1(I) procollagen mRNA. Sequencing of the amplified products in both cases identified a previously undescribed single base substitution.

5. Phenylketonuria: carrier detection at the population level by PCR.

Phenylketonuria (PKU) is an autosomal recessive disorder in which the defective gene encodes the enzyme phenylalanine hydroxylase (PAH). Single base substitutions have been identified in two mutant alleles that are common among Caucasians of northern European ancestry. These can be detected by hybridization with allele-specific oligonucleotides probes. DiLella, et al. [28] have shown that these alleles are easily detected in PCR-amplified fragments of genomic DNA that contain the sites of both mutations. Their study suggests that it is technically feasible to develop a program for population-based screening for carriers of the PKU alleles. Similar findings were reported in a study of another mutation at the PAH locus that is found in southern Europe and North Africa [29].

6. Cystic fibrosis and Huntington's disease: carrier screening and prenatal diagnosis by linkage analysis of PCR-amplified loci. *(Note added to the proof: It was announced on August 24 that the cystic fibrosis gene has been identified. If so, PCR will prove invaluable for identifying the spectrum of mutations affecting the gene , carrier screening , and prenatal diagnosis. At the investigative*

level, PCR will be valuable in studying the function of the gene product, which at present is unknown).

Cystic fibrosis and Huntington's disease are, respectively, autosomal recessive and autosomal dominant disorders, the genes for which have not yet been identified. For both diseases, as well as for a number of other single gene diseases for which the gene has been chromosomally mapped but not cloned, closely linked polymorphic loci have been identified that have been useful in carrier screening and prenatal diagnosis by linkage analysis of restriction fragment length polymorphisms in affected families. These methods have been simplified in a number of respects by the use of PCR. The need for Southern blotting and hybridization of genomic DNA can in many cases be obviated by PCR amplification of the polymorphic linked loci. In some cases, the PCR-amplified product is amenable to additional restriction enzymes, increasing the polymorphic content of the locus. Since less DNA is needed, buccal epithelial cells rather blood samples can be used [30].

In one report [31], eight pregnancies at risk for cystic fibrosis were monitored by first-trimester prenatal diagnosis. In all cases, linkage analysis by PCR produced results identical to conventional Southern blotting of genomic DNA.

Prenatal detection of a fetus at high risk for Huntington's disease has also been reported, using PCR amplification of a polymorphic locus linked to the HD locus [32].

7. <u>Identifying more than one defective gene in diseases previously thought to represent a single defect within a given population</u>. Many single gene diseases have been found to be caused by more than one defective allele (cf. example of ß-thal, above). However, it has sometimes appeared that only one particular defective allele is responsible for the disease in a given population. Because of the practical limitations of conventional cloning techniques, this assumption has been based on sequence data from only a few affected genes. PCR amplification has permitted many more individual alleles to be sequenced, and in the process, new mutations have been found within populations thought to have a single defective allele. Two examples of this follow.

Mitchell, et al. [33] studied the gene encoding ornithine ∂-aminotransferase, deficiency of which causes gyrate atrophy of the choroid and retina (GA), a rare monogenic disorders that occurs with increased frequency in the Finnish population. Although GA was previously thought to represent a single mutation in the Finnish population, PCR analysis of 16 Finnish GA pedigrees showed that the disease was caused by either of two separate missense mutations in the OAT gene . In a separate report, the same investigators describe a third mutation in OAT in Lebanese patients with GA [34].

In a similar study [35], Myerowitz examined the gene encoding the α chain of the lysosomal enzyme ß-N-acetylhexosaminidase A in Ashkenazi Jewish patients with classic Tay-Sachs disease. Before the advent of carrier-screening programs, this disorder occurred 10 times more frequently in this ethnic group than in the general population. Ashkenazi Jewish patients with classic Tay-Sachs disease have

appeared to be clinically and biochemically identical, and the usual assumption has been that they harbor the same α-chain mutation. However, in this study, only two of six genes from three patients studied by PCR had the same mutation as that found in the originally cloned patient cDNA, and only 30% of obligate heterozygotes carried the mutation. It thus appears that more than one mutation is responsible for the classic form of Tay-Sachs disease in the Ashkenazi Jewish population.

These findings may indicate that the original occurence of these mutations in these particular populations may not be simply the result of random events, but rather may be related to some selective advantage of the heterozygous state or some particular susceptibility of the underlying locus to mutation. It is of course also possible that the observed increased phenotype in these populations is in fact due to the random occurence of more than one mutation.

Some other single gene disease loci that have been studied by PCR are listed in Table XII.

Table XII. Some Additional Single-gene Disease Loci That Have Been Studied by PCR

Locus	Disease	Ref.
ß-globin	sickle cell disease	[1, 36-40]
dystrophin	Duchenne/Becker muscular dystrophy	[41- 43]
proinsulin	hyperinsulinemic mutant insulin syndrome	[44]
insulin receptor	insulin resistant diabetes mellitus	[45]
pyruvate dehydrogenase E1	lactic acidosis	[46]
aldolase B	hereditary fructose intolerance	[47]
ß-hexosaminidase	GM2 gangliosidosis (Sandhoff disease)	[48, 49]
liver/bone/kidney alkaline phosphatase	hypophosphatasia	[50]
α-1 antitrypsin	α-1-antitrypsin deficiency	[51, 52]
α-spectrin	hereditary elliptocytosis	[53]

glucose-6-phosphate dehydrogenase	G6PD deficiency	[54, 55]
serum cholinesterase	atypical cholinesterase	[56]
low density lipoprotein receptor	familial hypercholesterolemia	[16]
C1 INH	C1 esterase inhibitor deficiency	[21]
α-galactosidase	Fabry disease	[57]
HPRT	gout; Lesch-Nyhan syndrome	[58-61]
Apolipoprotein C-II	Apo CII-deficiency	[62]
leukocyte integrin ß chain	leukocyte adhesion deficiency	[63]

D. USE OF PCR IN THE DETECTION OF GENETIC POLYMORPHISM

In recent years, a great deal of effort has been expended in the construction of a human linkage map [64]. The goal has been to establish a panel of highly polymorphic markers detectable by RFLP at intervals of about 10 cM along each of the human chromosomes (1 cM = 1% recombination), so that no gene will be further away than 5 cM from a known polymorphic marker. A primary goal of this activity is the isolation of as yet unknown genes associated with a particular phenotype such as a disease by demonstrating genetic linkage to known loci, followed by application of recombiant DNA methods to obtain the target gene.

Most recent linkage analysis and other applications of highly polymorphic loci has been carried out by Southern blot-hybridization of genomic DNA, although some highly polymorphic loci, most notably in the HLA region, have been detectable by other methods such as serology. As previously mentioned, methods based on Southern blotting of genomic DNA have been constrained by the need for relatively large amounts of high quality DNA and by the relative insensitivity of RFLP analysis in detecting variations in DNA sequences. The advantages of PCR listed in Table I have been documented in several studies of human genetic polymorphism, principally in regard to the analysis of small amounts and/or degraded sources of DNA, and in the analysis of large numbers of individual samples at the DNA sequence level. Some examples follow.

1. Analysis of HLA-DQß sequences in autoimmune diseases. The polymorphic HLA-DQ alpha and beta loci encode the two chains of a class II HLA

molecule. The HLA-DQ alleles exist in very tight linkage disequilibrium with HLA-DR alleles, which are more readily detected serologically and have been found to be highly associated with a number of autoimmune diseases. Recent studies of DQ alleles at the DNA sequence level have suggested that some DR-associated diseases have an even stronger association with DQ. The advent of PCR has made it feasible to examine large numbers of DQ alleles at the DNA sequence level, either by direct DNA sequencing of amplified products, or, more conveniently, high stringency hybridization with allele-specific oligonucleotide probes. Two examples are discussed here, insulin-dependent diabetes mellitus (IDDM) and pemphigus vulgaris, of autoimmune diseases for which DQ associations have been sought by this method. Polymorphism of HLA-DP, another class II HLA product that has been difficult to study by classical methods, has also been characterized by PCR

Although IDDM is a multifactorial, polygenic disorder, there is a strong association with certain HLA alleles. It was recently reported that in a highly significant proportion of 39 patients with IDDM, both HLA-DQß alleles were from the group of alleles that contain a non-charged amino acid at position 57, whereas significantly fewer IDDM patients possessed any of the HLA-DQß alleles that have an aspartic acid residue at position 57 [65]. In an extension of this work [66], the HLA-DQß locus was amplified by PCR in 172 individuals in 27 families with multiple cases of IDDM, and the amplified DNA was hybridized with allele-specific oligonucleotides centered around codon 57. Sixty-five of 69 diabetic haplotypes were found to be homozygous for the non-Asp_{57} HLA-DQß alleles, vs. only 15 of 38 non-diabetic haplotypes. The relative risk for developing IDDM in non-Asp_{57} HLA-DQß homozygous individuals was 107, compared with Asp57 HLA-DQß heterozygous or homozygous individuals.

Recent work in Dr. J. D. Capra's laboratory here at Southwestern has been directed toward examining DQ alleles in a large population of patients with IDDM [67]. A total of 266 unrelated Caucasian IDDM patients and 203 normal Caucasian controls have been analyzed by PCR amplification of the relevant segment of the DQß locus. Some of the findings are summarized in Table IV. In contrast to the previous studies, these data from a very large population of IDDM patients show that a sizeable proportion (21%) of the patients possess at least one Asp_{57}-containing DQß allele. However, the finding that almost no IDDM patients are homozygous for Asp_{57}-containing DQß alleles was strikingly confirmed. Overall, the relative risk for IDDM in individuals possessing at least one non-Asp DQß allele was 42. The study also showed that some Asp_{57}-containing DQß alleles are more protective than others, and it suggested that other factors in addition to the DQß allele play a role in susceptibility to IDDM.

The ability to use PCR and allele-specific oligonucleotides to determine specific enotypes in large numbers of subjects, as illustrated by these investigations, should prove invaluable in unraveling the complicated genetics of diseases such as IDDM.

Table XIII. HLA-DQß Analysis in IDDM Patients and Controls

DQß genotype at position 57	Controls (n = 203)	IDDM (n = 266)	Rel. risk
Asp/Asp or Asp/blank	49	2	0.02
Asp/non-Asp	67	57	0.55
non-Asp/non-Asp non-Asp/blank	85	203	4.5
blank/blank	2	4	1.5

PCR analysis has also been applied to the association of pemphigus vulgaris with HLA-DR4 and HLA-DRw6 [68, 69]. Whereas 90% of pemphigus-associated DR4 haplotypes are of the subtype Dw10 and appear to be truly DR4-associated, the pemphigus association with HLA-DRw6 was found to reside in a rare DQ allele, DQß1.3.

2. Identifying individuals from DNA samples.

Several reports have documented that it is possible to identify PCR-amplified alleles in single leukocytes, buccal epithelial cells, hair roots, and sperm. In conjunction with several primer pairs that flank highly polymorphic loci suitable for analysis, simultaneous amplifications can be reliably analyzed from exceedingly small samples to allow identification of individuals by DNA "fingerprinting." Some examples of the descriptions of these techniques follow. The applications to forensic medicine and other legal situations such as paternity testing and immigration disputes, as well as more conventional medical situations such as the monitoring of bone marrow transplant engraftment, are obvious.

Amplification of DNA from single hairs. Higuchi et al. [70] successfully amplified HLA-DQα sequences from individual shed and freshly plucked hairs. In addition, genetically variable D-loop region mitochondrial DNA sequences were amplified from

individual roots and from a single hair shaft. The amplified DNA was successfully identified by fragment length differences, hybridization with allele-specific oligonucleotide probes, and direct DNA sequencing.

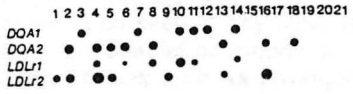
Hypervariable minisatellite DNA. Most vertebrate DNA contains minisatellite regions in which there occur large numbers of tandem repeats of a short sequence. In humans, two types of these regions have been useful for identifying DNA polymorphism. In both types, the polymorphism derives from the variable number of repeats that are found at a given locus, which in turn gives rise to varying length fragments when the DNA is digested with restriction enzymes that flank the region. In one type, the repeated sequence is specific for a particular locus, and such loci can be used as markers for linkage to a particular chromosomal region. In the second type, the repeated sequence is not specific for a particular locus, but occurs at several sites throughout the genome.

Jeffreys et al. [71], who have pioneered the use of the second type of minisatellite region in the identification of individual differences at the DNA level, have recently described the application of this technology to PCR amplified samples. The strategy adopted involves amplification of the entire minisatellite locus and hybridization of the PCR product with a specific probe. These investigators were routinely able to amplify segments of up to 6.6 kb, and to amplify simultaneously six different minisatellite loci from 0.01 µl samples of blood. The minisatellite loci used in this study are heterozygous in 66-97% of individuals, and the probability that two unrelated individuals would show the same pattern of 12 alleles was estimated at 2×10^{-5}. Successful amplification of multiple loci from individual lymphocytes, buccal epithelial cells, and hair roots was also demonstrated. Others have also reported the successful PCR-amplification of polymorphic loci containing tandem repeats [72, 73].

One important genetic difference to be determined is sex. The alphoid satellite family is a repetitive DNA family that is found only on the X and Y chromosomes. In one report in which dried blood specimens were the source of DNA, amplification of this locus in female DNA produced a 130-bp X-chromosome-specific fragment, whereas male DNA produced both a 170-bp Y-chromosome-specific and 130-bp X-chromosome-specific fragment [74]. Similarly, the putative testis-determining gene on the Y chromosome, *ZFY*, has been amplified by PCR [75]. Using these loci, it has been estimated that identification of fetal sex by the sixth week of gestation could be made by amplifying maternal peripheral blood DNA.

Amplication of DNA from single sperm. Li et al. [76] demonstrated that it is possible with PCR to amplify allelic sequences from individual sperm. They successfully co-amplified two unlinked genes from the sperm of an individual heterozygous at both loci, directly demonstrating Mendel's law of independent segregation (Table XIII, Figure 17). In addition to potential forensic applications, this technique could be applied for the construction of a more detailed human linkage map than is possible with the current method of pedigree analysis. In theory, sequential amplification and analysis of groups of three loci would provide a very accurate means

of ordering any number of tightly linked loci [77].



PCR analysis of individual sperm simultaneously amplified with *HLA DQA* and *LDLr* primers and analysed with allele-specific probes for the *DQA* alleles 1 and 2 (*DQA1*, *DQA2*) and *LDLr* alleles *1* and *2* (*LDLr1*, *LDLr2*). Water blanks are in positions 19-21. Examples of sperm where one allele at each locus is amplified are shown in positions 2, 5, 6, 8, 10, 11, 13, 14 and 16. The remaining samples can be divided into those that amplified more than one allele at a locus (position 4), only one allele at one locus (positions 1, 3, 7, 9, 12, 17, 18) or neither allele (position 15). During the first 20 cycles both loci were amplified simultaneously with both sets of primers present at 1 μM. A 2 μl aliquot from each reaction was then added to each of two tubes with 100 μl fresh PCR reaction buffer containing only one of the primer pairs. Forty-five additional PCR cycles were then performed.

Amplification of sequences at two different loci	
Total number of sperm examined	150
No signal	27
DQA1, *LDLr1*	21
DQA1, *LDLr2*	18
DQA2, *LDLr1*	14
DQA2, *LDLr2*	17
DQA1	14
DQA2	4
LDLr1	10
LDLr2	16
DQA1, *LDLr1*, *LDLr2*	2
DQA2, *LDLr1*, *LDLr2*	1
DQA1, *DQA2*	1
LDLr1, *LDLr2*	2
DQA1, *DQA2*, *LDLr1*, *LDLr2*	3
Controls	32
No signal	29
LDLr1	2
LDLr2	1

Figure 17 and Table XIV from reference 76. Sequences from two unlinked loci encoding HLA-DQA and the LDL receptor were simultaneously amplified in individual sperm from an individual heterozygous at both loci. Amplified DNA was hybridized with allele specific oligonucleotides. Instances in which one locus failed to amplify could represent either a technical problem or nondisjunction. Instances in which both alleles at one locus amplified could represent either a two sperm present or nondisjunction.

Amplification of mitochondrial DNA. Unlike nuclear DNA, mitochondrial DNA (mtDNA) is maternally inherited. One hundred and forty seven types of mtDNA havebeen identified in humans, and the interrelationship of these types is thought to be highly informative with regard to the sequence of evolution of human genetic diversity [78]. Amplification of mtDNA sequences by PCR should greatly facilitate the analysis of sequence diversity in much larger numbers of individuals than has been heretofore possible, and should lead to significant refinement of the human phylogenetic tree of mtDNA types [78].

Amplification of ancient DNA. A series of reports by Pääbo, et al. [79-82] have discussed the possibilities and limitations of amplifying sequences from ancient DNA such as that found under certain conditions in archeologic specimens. The DNA obtained from such specimens has invariably been of low average molecular size and damaged by oxidative processes, such that it has not been feasible to use conventional cloning methods to isolate specific segments. However, by PCR amplification, several short sequences of potential interest have been obtained. Mitochondrial DNA, which is present in multiple copies per cell, has been the most easily amplified. Amplification of a polymorphic stretch of region V of mtDNA from a 7000 year old brain excavated in Florida revealed a sequence that differs from any

previously detected among populations descended from pre-Columbian North American inhabitants [79]. Similarly, amplification of liver DNA from a 4000 year old Egyptian mummy yielded a sequence from region V that was identical to a known sequence, whereas a sequence from the D loop carried a nucleotide at one position that has not been seen in any other human sequence determined to date [82]. In addition to its potential anthropologic applications, such analysis may be useful in evolutionary studies such as phylogenetic assignment of extinct species for which specimens are available.

III. APPLICATION OF PCR TO CELL AND MOLECULAR BIOLOGY AND BIOTECHNOLOGY

PCR has been successfully applied to the isolation of new members of gene families for in which known members contain regions of conserved sequence. In these studies, sets of primers are used containing alternative nucleotides at positions of variability among the known sequences, and the amplified products individually cloned and sequenced. Examples of this approach include the recent isolation of new members of the G-protein-coupled receptor family from cDNA [83], of the potassium channel family from genomic DNA [84], and of the tyrosine kinase families from cDNA [85] and genomic DNA [84] . This method should lead to the rapid identification of a large number of new genes and gene products. An alternative to the construction of degenerate primer sets is the use of inosine for ambiguous nucleotide positions, since this purine nucleotide forms stable base pairs with both purines and pyrimidines [86, 87].

Significant advances are likely to come from the use of PCR to analyze the genes and and gene products of small numbers of cells or single cells. This will permit biochemical analysis of processes such as growth control that are often regulated by small numbers of cells expressing pathways that involve mRNA transcripts that are short-lived and have a low copy number. One recent example of such an application was reported by Rappolee, et al. [88, 89], who used PCR to analyze the mRNA products of macrophages isolated directly from wound sites. These cells were found to express transcripts for transforming growth factors-α and -ß, platelet-derived growth factor, and insulin-like growth factor-1, all factors thought to mediate various stages of wound healing but not previously demonstrated as products of wound macrophages. Similar studies of mRNA from small numbers of cells should allow detailed analysis of factors that lead to alternative splicing of mRNA under varying conditions [41, 54, 90, 91].

A host of new methods based on PCR have been reported and are beginning to revolutionize the conduct of research in molecular biology. Among the applications reported are new methods for cloning full-length cDNA [92], screening for mutations [93], in vitro mutagenesis [94], engineering hybrid genes [95], gene mapping [96, 97], and DNA sequencing [98-101]. PCR-based methods are likely tohave a cumulative effect of accelerating the rate of application of molecular biology to solving clinical problems.

PCR-based technology is also likely to have an impact of the production of biological therapeutic agents. For example, it has been shown that immunoglobulin heavy and light chain variable domains can be amplified by PCR and then cloned into suitable vectors for sequencing and expression [102, 103]. With PCR-based methods for the rapid construction of hybrid DNA molecules [94, 95], it should be possible to construct antibodies of defined specificity that have complementarity-determining regions from appropriate mouse hybridomas, but human variable region framework constant region sequences. Such antibodies would be less likely than conventional mouse monoclonal antibodies to be targets of an immune response that might limit their therapeutic effectiveness or safety.

IV. CONCLUSION

Most of the early uses of PCR in clinical medicine have been in areas in which molecular biology had already provided a large number of tools. These are listed in Table XV, and have formed the basis for most of this Grand Rounds.

Table XV. Applications of PCR to Clinical Medicine - Early Successes

Detection of virus in ambiguous cases (HIV, HBV, HTLV-II, others)
Identification of virus in patients with neurologic disease
Identification of papillomavirus subtype associated with cervical CA
Detection of minimal residual disease in hematologic malignancies
Identification of *ras* mutations in pancreatic, thyroid, colon CA; AML, ALL, MDS
Identification of heterogeneity in Ph' chromosome in CML and ALL
Identification of HLA-DQß alleles associated with IDDM
Characterization of new mutations in numerous single gene diseases
Carrier screening and prenatal detection of several single gene diseases

Table XVI contains some predictions for future applications of PCR-based technology to clinical medicine, some of which may provide future topics for these Grand Rounds.

Table XVI. Applications of PCR to Clinical Medicine - Predictions for the Future

Identification of viruses in chronic diseases with a viral pathogenesis
Rapid diagnosis of acute infectious disease
Extensive characterization of the genetic alterations in malignancy
Characterization of the antibody and T cell receptor sequences in autoimmunity
Widespread screening for genetic disease in single gene and polygenic disorders
Widespread use of DNA fingerprinting
Molecular engineering of therapeutically useful products
Insights into pathogenesis and treatment from studies of single cells
Complementation of human genome sequencing project

References

1. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA , Arnheim N. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350-4 (1985).

2. Scharf SJ, Horn GT , Erlich HA. Direct cloning and sequence analysis of enzymatically amplified genomic sequences. Science 233: 1076-8 (1986).

3. Saiki RK, Bugawan TL, Horn GT, Mullis KB , Erlich HA. Analysis of enzymatically amplified beta-globin and HLA-DQ alpha DNA with allele-specific oligonucleotide probes. Nature 324: 163-6 (1986).

4. Orkin SH. Molecular genetics and inherited human disease. The Metabolic Basis of Inherited Disease. Scriver CR, Beaudet AL, Sly WS and Valle D ed. 1989 McGraw-Hill. NY.

5. Cai SP, Zhang JZ, Huang DH, Wang ZX , Kan YW. A simple approach to prenatal diagnosis of beta-thalassemia in a geographic area where multiple mutations occur. Blood 71: 1357-60 (1988).

6. Cai SP, Chang CA, Zhang JZ, Saiki RK, Erlich HA , Kan YW. Rapid prenatal diagnosis of beta thalassemia using DNA amplification and nonradioactive probes. Blood 73: 372-4 (1989).

7. Antonarakis SE, Kang J, Lam VM, Tam JW, Li AM. Molecular characterization

of beta-globin gene mutations in patients with beta-thalassaemia intermedia in south China. Br J Haematol 70: 357-61 (1988).

8. Amselem S, Nunes V, Vidaud M, Estivill X, Wong C, dAuriol L, Vidaud D, Galibert F, Baiget M, Goossens M. Determination of the spectrum of beta-thalassemia genes in Spain by use of dot-blot analysis of amplified beta-globin DNA. Am J Hum Genet 43: 95-100 (1988).

9. Wong C, Dowling CE, Saiki RK, Higuchi RG, Erlich HA , Kazazian HHJ. Characterization of beta-thalassaemia mutations using direct genomic sequencing of amplified single copy DNA. Nature 330: 384-6 (1987).

10. Chan V, Chan TK, Kan YW , Todd D. A novel beta-thalassemia frameshift mutation (codon 14/15), detectable by direct visualization of abnormal restriction fragment in amplified genomic DNA. Blood 72: 1420-3 (1988).

11. Thein SL, Hesketh C, Brown JM, Anstey AV , Weatherall DJ. Molecular characterization of a high A2 beta thalassemia by direct sequencing of single strand enriched amplified genomic DNA. Blood 73: 924-30 (1989).

12. Fucharoen S, Fucharoen G, Fucharoen P , Fukumaki Y. A novel ochre mutation in the beta-thalassemia gene of a Thai. Identification by direct cloning of the entire beta-globin gene amplified using polymerase chain reactions. J Biol Chem 264: 7780-3 (1989).

13. Spitzer SG, Pendurthi UR, Kasper CK , Bajaj SP. Molecular defect in factor IXBm Lake Elsinore. Substitution of Ala390 by Val in the catalytic domain. J Biol Chem 263: 10545-8 (1988).

14. Tsang TC, Bentley DR, Mibashan RS , Giannelli F. A factor IX mutation, verified by direct genomic sequencing, causes haemophilia B by a novel mechanism. Embo J 7: 3009-15 (1988).

15. Denton PH, Fowlkes DM, Lord ST , Reisner HM. Hemophilia B Durham: a mutation in the first EGF-like domain of factor IX that is characterized by polymerase chain reaction. Blood 72: 1407-11 (1988).

16. Leitersdorf E, Hobbs HH, Fourie AM, Jacobs M, van der Westhuyzen DR. Deletion in the first cysteine-rich repeat of low density lipoprotein receptor impairs its transport but not lipoprotein binding in fibroblasts from a subject with familial hypercholesterolemia. Proc Natl Acad Sci U S A 85: 7912-6 (1988).

17. Winship PR, Rees DJ , Alkan M. Detection of polymorphisms at cytosine phosphoguanadine dinucleotides and diagnosis of haemophilia B carriers. Lancet 1:

631-4 (1989).

18. Graham JB, Kunkel GR, Tennyson GS, Lord ST , Fowlkes DM. The Malmo polymorphism of factor IX: establishing the genotypes by rapid analysis of DNA. Blood 73: 2104-7 (1989).

19. Zhang M, Chen SH, Scott CR , Thompson AR. The factor IX BamHI polymorphism: T-to-G transversion at the nucleotide sequence -561. The BamHI/MSPI haplotypes in blacks and Caucasians. Hum Genet 82: 283-4 (1989).

20. Inaba H, Fujimaki M, Kazazian HHJ , Antonarakis SE. Mild hemophilia A resulting from Arg-to-Leu substitution in exon 26 of the factor VIII gene. Hum Genet 81: 335-8 (1989).

21. Arai M, Inaba H, Higuchi M, Antonarakis SE, Kazazian HHJ, Fujimaki M , Hoyer LW. Direct characterization of factor VIII in plasma: detection of a mutation altering a thrombin cleavage site (arginine-372----histidine). Proc Natl Acad Sci U S A 86: 4277-81 (1989).

22. Ginsburg D, Konkle BA, Gill JC, Montgomery RR, Bockenstedt PL, Johnson TA , Yang AY. Molecular basis of human von Willebrand disease: analysis of platelet von Willebrand factor mRNA. Proc Natl Acad Sci U S A 86: 3723-7 (1989).

23. Newman PJ, Gorski J, White GC2, Gidwitz S, Cretney CJ , Aster RH. Enzymatic amplification of platelet-specific messenger RNA using the polymerase chain reaction. J Clin Invest 82: 739-43 (1988).

24. Newman PJ, Derbes RS , Aster RH. The human platelet alloantigens, PIA1 and PIA2, are associated with a leucine33/proline33 amino acid polymorphism in membrane glycoprotein IIIa, and are distinguishable by DNA typing. J Clin Invest 83: 1778-81 (1989).

25. Lee B, Vissing H, Ramirez F, Rogers D , Rimoin D. Identification of the molecular defect in a family with spondyloepiphyseal dysplasia. Science 244: 978-80 (1989).

26. Labhard ME, Wirtz MK, Pope FM, Nicholls AC , Hollister DW. A cysteine for glycine substitution at position 1017 in an alpha 1(I) chain of type I collagen in a patient with mild dominantly inherited osteogenesis imperfecta. Mol Biol Med 5: 197-207 (1988).

27. Patterson E, Smiley E , Bonadio J. RNA sequence analysis of a perinatal lethal osteogenesis imperfecta mutation. J Biol Chem 264: 10083-7 (1989).

28. DiLella AG, Huang WM , Woo SL. Screening for phenylketonuria mutations by DNA amplification with the polymerase chain reaction. Lancet 1: 497-9 (1988).

29. Lyonnet S, Caillaud C, Rey F, Berthelon M, Frezal J, Rey J , Munnich A. Molecular genetics of phenylketonuria in Mediterranean countries: a mutation associated with partial phenylalanine hydroxylase deficiency. Am J Hum Genet 44: 511-7 (1989).

30. Lench N, Stanier P , Williamson R. Simple non-invasive method to obtain DNA for gene analysis. Lancet 1: 1356-8 (1988).

31. Gasparini P, Novelli G, Savoia A, Dallapiccola B , Pignatti PF. First-trimester prenatal diagnosis of cystic fibrosis using the polymerase chain reaction: report of eight cases. Prenat Diagn 9: 349-55 (1989).

32. McIntosh I, Curtis A, Millan FA , Brock DJ. Prenatal exclusion testing for Huntington disease using the polymerase chain reaction. Am J Med Genet 32: 274-6 (1989).

33. Mitchell GA, Brody LC, Sipila I, Looney JE, Wong C, Engelhardt JF, Patel AS, Steel G, Obie C, Kaiser KM , et al. At least two mutant alleles of ornithine delta-aminotransferase cause gyrate atrophy of the choroid and retina in Finns. Proc Natl Acad Sci U S A 86: 197-201 (1989).

34. Mitchell GA, Brody LC, Looney J, Steel G, Suchanek M, Dowling C, Der KV, Kaiser KM , Valle D. An initiator codon mutation in ornithine-delta-aminotransferase causing gyrate atrophy of the choroid and retina. J Clin Invest 81: 630-3 (1988).

35. Myerowitz R. Splice junction mutation in some Ashkenazi Jews with Tay-Sachs disease: evidence against a single defect within this ethnic group. Proc Natl Acad Sci U S A 85: 3955-9 (1988).

36. Embury SH, Scharf SJ, Saiki RK, Gholson MA, Golbus M, Arnheim N , Erlich HA. Rapid prenatal diagnosis of sickle cell anemia by a new method of DNA analysis. N Engl J Med 316: 656-61 (1987).

37. Jinks DC, Minter M, Tarver DA, Vanderford M, Hejtmancik JF , McCabe ER. Molecular genetic diagnosis of sickle cell disease using dried blood specimens on blotters used for newborn screening. Hum Genet 81: 363-6 (1989).

38. Kulozik AE, Lyons J, Kohne E, Bartram CR , Kleihauer E. Rapid and non-radioactive prenatal diagnosis of beta thalassaemia and sickle cell disease: application of the polymerase chain reaction (PCR). Br J Haematol 70: 455-8 (1988).

39. Saiki RK, Chang CA, Levenson CH, Warren TC, Boehm CD, Kazazian HHJ , Erlich HA. Diagnosis of sickle cell anemia and beta-thalassemia with enzymatically amplified DNA and nonradioactive allele-specific oligonucleotide probes. N Engl J Med 319: 537-41 (1988).

40. Wu DY, Ugozzoli L, Pal BK , Wallace RB. Allele-specific enzymatic amplification of beta-globin genomic DNA for diagnosis of sickle cell anemia. Proc Natl Acad Sci U S A 86: 2757-60 (1989).

41. Feener CA, Koenig M , Kunkel LM. Alternative splicing of human dystrophin mRNA generates isoforms at the carboxy terminus. Nature 338: 509-11 (1989).

42. Chelly J, Kaplan JC, Maire P, Gautron S , Kahn A. Transcription of the dystrophin gene in human muscle and non-muscle tissue. Nature 333: 858-60 (1988).

43. Chamberlain JS, Gibbs RA, Ranier JE, Nguyen PN , Caskey CT. Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. Nucleic Acids Res 16: 11141-56 (1988).

44. Miyano M, Nanjo K, Chan SJ, Sanke T, Kondo M , Steiner DF. Use of in vitro DNA amplification to screen family members for an insulin gene mutation. Diabetes 37: 862-6 (1988).

45. Kobayashi M, Sasaoka T, Takata Y, Ishibashi O, Sugibayashi M, Shigeta Y, Hisatomi A, Nakamura E, Tamaki M , Teraoka H. Insulin resistance by unprocessed insulin proreceptors point mutation at the cleavage site. Biochem Biophys Res Commun 153: 657-63 (1988).

46. Endo H, Hasegawa K, Narisawa K, Tada K, Kagawa Y , Ohta S. Defective gene in lactic acidosis: abnormal pyruvate dehydrogenase E1 alpha-subunit caused by a frame shift. Am J Hum Genet 44: 358-64 (1989).

47. Cross NC, Tolan DR , Cox TM. Catalytic deficiency of human aldolase B in hereditary fructose intolerance caused by a common missense mutation. Cell 53: 881-5 (1988).

48. Paw BH, Kaback MM , Neufeld EF. Molecular basis of adult-onset and chronic GM2 gangliosidoses in patients of Ashkenazi Jewish origin: substitution of serine for glycine at position 269 of the alpha-subunit of beta-hexosaminidase. Proc Natl Acad Sci U S A 86: 2413-7 (1989).

49. Nakano T , Suzuki K. Genetic cause of a juvenile form of Sandhoff disease. Abnormal splicing of beta-hexosaminidase beta chain gene transcript due to a point

mutation within intron 12. J Biol Chem 264: 5155-8 (1989).

50. Weiss MJ, Cole DE, Ray K, Whyte MP, Lafferty MA, Mulivor RA , Harris H. A missense mutation in the human liver/bone/kidney alkaline phosphatase gene causing a lethal form of hypophosphatasia. Proc Natl Acad Sci U S A 85: 7666-9 (1988).

51. Newton CR, Kalsheker N, Graham A, Powell S, Gammack A, Riley J , Markham AF. Diagnosis of alpha 1-antitrypsin deficiency by enzymatic amplification of human genomic DNA and direct sequencing of polymerase chain reaction products. Nucleic Acids Res 16: 8233-43 (1988).

52. Hejtmancik JF, Holcomb JD, Howard J , Vanderford M. In vitro amplification of the alpha 1-antitrypsin gene: application to prenatal diagnosis. Prenat Diagn 9: 177-86 (1989).

53. Roux AF, Morle F, Guetarni D, Colonna P, Sahr K, Forget BG, Delaunay J , Godet J. Molecular basis of Sp alpha I/65 hereditary elliptocytosis in North Africa: insertion of a TTG triplet between codons 147 and 149 in the alpha-spectrin gene from five unrelated families. Blood 73: 2196-201 (1989).

54. Hirono A, Beutler E. Alternative splicing of human glucose-6-phosphate dehydrogenase messenger RNA in different tissues. J Clin Invest 83: 343-6 (1989).

55. Hirono A, Beutler E. Molecular cloning and nucleotide sequence of cDNA for human glucose-6-phosphate dehydrogenase variant A(-). Proc Natl Acad Sci U S A 85: 3951-4 (1988).

56. McGuire MC, Nogueira CP, Bartels CF, Lightstone H, Hajra A, Van dSAF, Lockridge O , La DBN. Identification of the structural mutation responsible for the dibucaine-resistant (atypical) variant form of human serum cholinesterase. Proc Natl Acad Sci U S A 86: 953-7 (1989).

57. Bernstein HS, Bishop DF, Astrin KH, Kornreich R, Eng CM, Sakuraba H , Desnick RJ. Fabry disease: six gene rearrangements and an exonic point mutation in the alpha-galactosidase gene. J Clin Invest 83: 1390-9 (1989).

58. Vrieling H, Simons JW , van Zeeland AA. Nucleotide sequence determination of point mutations at the mouse HPRT locus using in vitro amplification of HPRT mRNA sequences. Mutat Res 198: 107-13 (1988).

59. Cariello NF, Scott JK, Kat AG, Thilly WG , Keohavong P. Resolution of a missense mutant in human genomic DNA by denaturing gradient gel electrophoresis and direct sequencing using in vitro DNA amplification: HPRT Munich. Am J Hum

Genet 42: 726-34 (1988).

60. Davidson BL, Chin SJ, Wilson JM, Kelley WN , Palella TD. Hypoxanthine-guanine phosphoribosyltransferase. Genetic evidence for identical mutations in two partially deficient subjects. J Clin Invest 82: 2164-7 (1988).

61. Simpson D, Crosby RM , Skopek TR. A method for specific cloning and sequencing of human hprt cDNA for mutation analysis. Biochem Biophys Res Commun 151: 487-92 (1988).

62. Fojo SS, Stalenhoef AF, Marr K, Gregg RE, Ross RS , Brewer HBJ. A deletion mutation in the ApoC-II gene (ApoC-II Nijmegen) of a patient with a deficiency of apolipoprotein C-II. J Biol Chem 263: 17913-6 (1988).

63. Kishimoto TK, OConner K , Springer TA. Leukocyte adhesion deficiency. Aberrant splicing of a conserved integrin sequence causes a moderate deficiency phenotype. J Biol Chem 264: 3588-95 (1989).

64. White RL, Lalouel J-M. Linked sets of genetic markers for human chromosomes. Ann. Rev. Genet. 22: 259-280 (1988).

65. Todd JA, Bell JI , McDevitt HO. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 329: 599-604 (1987).

66. Morel PA, Dorman JS, Todd JA, McDevitt HO , Trucco M. Aspartic acid at position 57 of the HLA-DQ beta chain protects against type I diabetes: a family study. Proc Natl Acad Sci U S A 85: 8111-5 (1988).

67. Baisch J, Weeks T, Gilles R, Hoover M, Stastny P , Capra JD. Molecular analysis of DQ genotypes to assess disease susceptibility in insulin dependent diabetes mellitus. A study of 266 patients. Submitted for publication

68. Scharf SJ, Friedmann A, Brautbar C, Szafer F, Steinman L, Horn G, Gyllensten U , Erlich HA. HLA class II allelic variation and susceptibility to pemphigus vulgaris. Proc Natl Acad Sci U S A 85: 3504-8 (1988).

69. Scharf SJ, Long CM , Erlich HA. Sequence analysis of the HLA-DR beta and HLA-DQ beta loci from three Pemphigus vulgaris patients. Hum Immunol 22: 61-9 (1988).

70. Higuchi R, von Beroldingen, Sensabaugh GF , Erlich HA. DNA typing from single hairs. Nature 332: 543-6 (1988).

71. Jeffreys AJ, Wilson V, Neumann R , Keyte J. Amplification of human

minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. Nucleic Acids Res 16: 10953-71 (1988).

72. Horn GT, Richards B , Klinger KW. Amplification of a highly polymorphic VNTR segment by the polymerase chain reaction. Nucleic Acids Res 17: 2140 (1989).

73. Weber JL , May PE. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet 44: 388-96 (1989).

74. Witt M , Erickson RP. A rapid method for detection of Y-chromosomal DNA from dried blood specimens by the polymerase chain reaction. Hum Genet 82: 271-4 (1989).

75. Ebensperger C, Studer R , Epplen JT. Specific amplification of the ZFY gene to screen sex in man. Hum Genet 82: 289-90 (1989).

76. Li HH, Gyllensten UB, Cui XF, Saiki RK, Erlich HA , Arnheim N. Amplification and analysis of DNA sequences in single human sperm and diploid cells. Nature 335: 414-7 (1988).

77. Boehnke M, Arnheim N, Li H , Collins FS. Fine-structure genetic mapping of human chromosomes using the polymerase chain reaction on single sperm: experimental design considerations. Am J Hum Genet 45: 21-32 (1989).

78. Wrischnik LA, Higuchi RG, Stoneking M, Erlich HA, Arnheim N , Wilson AC. Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. Nucleic Acids Res 15: 529-42 (1987).

79. Pääbo S, Gifford JA , Wilson AC. Mitochondrial DNA sequences from a 7000-year old brain. Nucleic Acids Res 16: 9775-87 (1988).

80. Pääbo S , Wilson AC. Polymerase chain reaction reveals cloning artefacts [letter]. Nature 334: 387-8 (1988).

81. Pääbo S, Higuchi RG , Wilson AC. Ancient DNA and the polymerase chain reaction. The emerging field of molecular archaeology. J Biol Chem 264: 9709-12 (1989).

82. Pääbo S. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci U S A 86: 1939-43 (1989).

83. Libert F, Parmentier M, Lefort A, Dinsart C, Van Sande J, Maenhaut C, Simons MJ, Dumont JE , Vassart G. Selective amplification and cloning of four new members of the G protein-coupled receptor family. Science 244: 569-72 (1989).

84. Kamb A, Weir M, Rudy B, Varmus H , Kenyon C. Identification of genes from pattern formation, tyrosine kinase, and potassium channel families by DNA amplification. Proc Natl Acad Sci U S A 86: 4372-6 (1989).

85. Wilks AF. Two putative protein-tyrosine kinases identified by application of the polymerase chain reaction. Proc Natl Acad Sci U S A 86: 1603-7 (1989).

86. Ehlen T , Dubeau L. Detection of ras point mutations by polymerase chain reaction using mutation-specific, inosine-containing oligonucleotide primers. Biochem Biophys Res Commun 160: 441-7 (1989).

87. Knoth K, Roberds S, Poteet C , Tamkun M. Highly degenerate, inosine-containing primers specifically amplify rare cDNA using the polymerase chain reaction. Nucleic Acids Res 16: 10932 (1988).

88. Rappolee DA, Mark D, Banda MJ , Werb Z. Wound macrophages express TGF-alpha and other growth factors in vivo: analysis by mRNA phenotyping. Science 241: 708-12 (1988).

89. Rappolee DA, Wang A, Mark D , Werb Z. Novel method for studying mRNA phenotypes in single or small numbers of cells. J Cell Biochem 39: 1-11 (1989).

90. Seino S , Bell GI. Alternative splicing of human insulin receptor messenger RNA. Biochem Biophys Res Commun 159: 312-6 (1989).

91. Conboy JG, Chan J, Mohandas N , Kan YW. Multiple protein 4.1 isoforms produced by alternative splicing in human erythroid cells. Proc Natl Acad Sci U S A 85: 9062-5 (1988).

92. Frohman MA, Dush MK , Martin GR. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc Natl Acad Sci U S A 85: 8998-9002 (1988).

93. Sheffield VC, Cox DR, Lerman LS , Myers RM. Attachment of a 40-base-pair G + C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. Proc Natl Acad Sci U S A 86: 232-6 (1989).

94. Ho SN, Hunt HD, Horton RM, Pullen JK , Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77: 51-59 (1989).

95. Horton RM, Hunt HD, Ho SN, Pullen JK , Pease LR. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. Gene 77:

61-68 (1989).

96. Loh EY, Elliott JF, Cwirla S, Lanier LL , Davis MM. Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. Science 243: 217-20 (1989).

97. Ochman H, Gerber AS , Hartl DL. Genetic applications of an inverse polymerase chain reaction. Genetics 120: 621-3 (1988).

98. Gyllensten 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 U S A 85: 7652-6 (1988).

99. Stoflet ES, Koeberl DD, Sarkar G , Sommer SS. Genomic amplification with transcript sequencing. Science 239: 491-4 (1988).

100. Sarkar G , Sommer SS. Access to a messenger RNA sequence or its protein product is not limited by tissue or species specificity. Science 244: 331-4 (1989).

101. Gyllensten UB. PCR and DNA sequencing. BioTechniques 7: 700-708 (1989).

102. Orlandi R, Gussow DH, Jones PT , Winter G. Cloning immunoglobulin variable domains for expression by the polymerase chain reaction. Proc Natl Acad Sci U S A 86: 3833-7 (1989).

103. Larrick JW, Danielsson L, Brenner CA, Abrahamson M, Fry KE , Borrebaeck CA. Rapid cloning of rearranged immunoglobulin genes from human hybridoma cells using mixed primers and the polymerase chain reaction. Biochem Biophys Res Commun 160: 1250-6 (1989).