Nucleic Acid Amplification In Vitro: Detection of Sequences with Low Copy Numbers and Application to Diagnosis of Human Immunodeficiency Virus Type 1 Infection

JOHN C. GUATELLI,[1] THOMAS R. GINGERAS,[2] AND DOUGLAS D. RICHMAN[1]*

Departments of Medicine and Pathology, University of California-San Diego School of Medicine, and San Diego Veterans Administration Medical Center, San Diego, California 92161,[1] and The Salk Institute Biotechnology/Industrial Associates, Inc., La Jolla, California 92037[2]

INTRODUCTION

There are two important factors for detecting low-copy-number, blood-borne viral pathogens. The genomic copy number estimated to be present in 1 ml of blood of hepatitis B virus, cytomegalovirus, Epstein-Barr virus, and human immunodeficiency virus type 1 (HIV-1) ranges from $\leq 10^{12}$ deoxyribonucleic acid (DNA) copies for hepatitis B virus to $\leq 10^6$ ribonucleic acid (RNA) copies or $\leq 10^3$ DNA copies for HIV-1 (26); values for Epstein-Barr virus and cytomegalovirus are $\leq 10^4$ and $\leq 10^6$ copies of DNA, respectively. Nucleic acid hybridization reactions, performed with highly ^{32}P-labeled nucleic acid probes, are the most sensitive in vitro method for detecting such low-copy-number viral targets. Such probes (called "nick-translated" probes because of the incorporation of many ^{32}P-linked nucleotides, used to repair enzymatically "nicked" DNA probes) are capable of detecting 10^4 to 10^5 copies of complementary target nucleic

acid. Alternatively, if only one ^{32}P label is attached (usually at the 5' or 3' end) to the probe, the sensitivity of this detection system decreases to 10^6 to 10^7 copies. Nucleic acid hybridizations performed with a nonisotopic detection system (for example, using a colorimetric signal generated by alkaline phosphatase) exhibit sensitivities of detection which range from 10^5 to 10^7 copies of complementary target molecules. In general, the nonisotopic detection systems are 10- to 10^3-fold less sensitive than ^{32}P-labeled probes. The requirement of detecting a very low number of target molecules is thus made more difficult by the limited sensitivities of the existing detection systems.

Only two alternatives are available to meet the challenge of detecting low-copy-number viral targets. Either significantly more sensitive detection systems need to be developed, or the original copy number of the target virus needs to be elevated before a detection assay is initiated. Recently, the most spectacular results in detecting low-copy-number
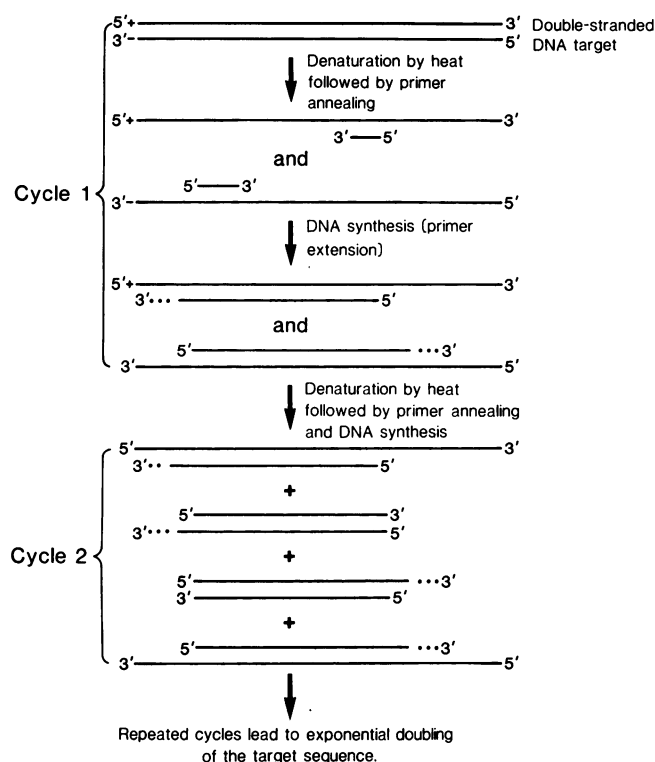
FIG. 1. Steps required for amplification of a DNA sequence by PCR. Two oligonucleotide primers are utilized. These primers are complementary to sequences located several hundred base pairs apart on opposite strands of the target. A cycle of PCR consists of denaturation of the target in the presence of a large excess of primers, followed by primer annealing and extension. The product molecules can serve as templates in the following cycles. Repeated cycles lead to exponential accumulation of the target region bracketed by the two primers.

target molecules have been achieved by using the latter strategy. This review describes the development of a nucleic acid amplification system in vitro and how the detection and diagnosis of HIV-1 have served as a model system for the application of nucleic acid amplification to the detection of viruses.

METHODOLOGY OF SEQUENCE AMPLIFICATION BY PCR

The Method

The first strategy to be developed for nucleic acid target amplification in vitro has been termed the polymerase chain reaction (PCR) (22). PCR can increase the copy number of specific nucleic acid sequences of 100 to 2,000 base pairs in length by 10^6-fold in an in vitro reaction of only a few hours in duration (Fig. 1). This reaction requires two oligonucleotide primers which are complementary to target sequences several hundred base pairs apart from each other. Each oligonucleotide is complementary to opposite DNA strands of the target and subsequent amplification products. For a DNA target, the first step involves thermal denaturation of the double-stranded target molecules in the presence of a large molar excess of primers. The primers are then annealed to their complementary target sequences by temperature reduction. Next, a DNA polymerase is added, catalyzing a

primer-directed DNA synthesis reaction. The result is an approximate doubling of the amount of target sequence in the sample. Extension from each primer proceeds in opposite, overlapping directions. Thus, DNA synthesis primed from one oligonucleotide generates a new strand with a region complementary to the other oligonucleotide primer. The product strands can then serve as templates in successive DNA synthesis reactions. By repeating the cycle of denaturation, primer annealing, and DNA synthesis (primer extension), the copy number of target sequence is increased exponentially (22). The theoretical amplification can be expressed as 2^n, where n is the number of cycles performed. In practice, the extent of amplification is $(1 + x)^n$, where x is the average efficiency of each cycle (2, 32). Thus, with an average per-cycle efficiency of 67%, the increase in copy number per target molecule is 30,000-fold after 20 cycles of PCR.

Several features of the PCR process deserve emphasis. First, the products of this reaction are double-stranded DNA molecules of specific molecular weight which correspond to the portion of the target bracketed by the two primers. As a consequence, the primers themselves become the ends of the product molecules. Oligonucleotide primers with 5' extensions not complementary to the target still function effectively as primers (22, 33). These extensions will become built into the product molecules as the reaction proceeds (Fig. 2). This is an important feature because it allows any desired sequence to be appended to one or both ends of the product fragment. For example, restriction endonuclease sites may be appended to an amplified fragment to facilitate its subsequent cloning (33).

Second, RNA molecules may also serve as targets for PCR amplification. The use of RNA sequences as targets enables the detection of RNA viruses. It also allows the detection of virus-specific RNA transcripts from both RNA and DNA viruses, enabling an assessment of the transcriptional activity of the pathogen. To amplify an RNA sequence, the first primer-directed DNA synthesis reaction is performed with RNA-dependent reverse transcriptase. A DNA-dependent polymerase is added to synthesize the second strand of complementary DNA, and multiple cycles of a typical PCR reaction are then performed (2, 15). The amplification products are again double-stranded DNA fragments.

Third, the utility and efficacy of the PCR amplification scheme have been increased by using the DNA polymerase of the thermophilic bacterium *Thermus aquaticus* (*Taq*). The *Taq* DNA polymerase is thermostable and thus able to tolerate the repeated thermal denaturation steps of PCR without significant loss of activity (4, 6, 16, 31). The original PCR reactions were performed with the Klenow fragment of *Escherichia coli* DNA polymerase I. Because Klenow DNA polymerase is temperature sensitive, it was necessary to add fresh enzyme before the DNA-synthetic step of each cycle. In contrast, once *Taq* DNA polymerase is added to a reaction mixture of target, deoxyribonucleotides, and oligonucleotide primers, the PCR process requires only repeated cycles of incubations at the temperatures necessary for target denaturation, primer annealing, and DNA synthesis. These cycles can be accomplished by using a computer-controlled heat block (thermal cycler). Thus, the utilization of *Taq* polymerase has made it possible to develop an automated sequence amplification process.

The *Taq* polymerase also enables PCR to be performed with increased efficiency and specificity (6, 31). These properties are most likely attributed to the elevated annealing

+5'————————————————3' Double-stranded
-3'————————————————5' DNA target

A PCR cycle using primers
with non-complementary
5' extensions

+5'————————————————3'

and

-3'————————————————5'

A second cycle of PCR

5'————————————————3'

3'————————————————5'

A third cycle of PCR
(only short products shown)

continued amplification of target molecules
now containing the desired appended sequence
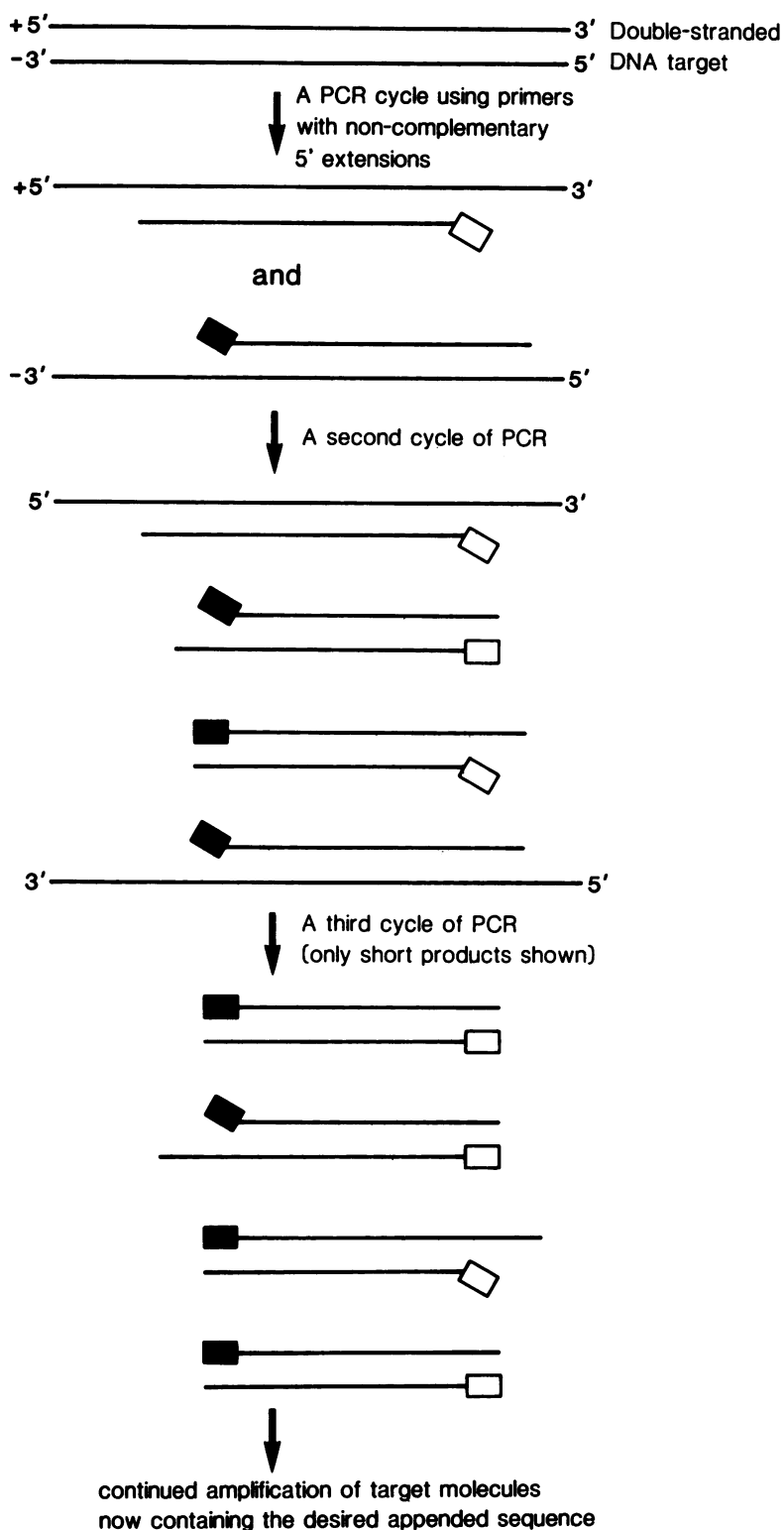
FIG. 2. Mechanism by which any desired sequence can be incorporated into the ends of PCR amplification products. Primers are utilized which have non-target-complementary 5' extensions. These extensions become built into the ends of the double-stranded product molecules as the PCR process proceeds. Symbols: □, noncomplementary 5' sequence on antisense oligonucleotide primers; ■, noncomplementary 5' sequence on sense oligonucleotide primers.

temperatures possible with *Taq*, as discussed below. Compared with Klenow, *Taq* polymerase can effectively amplify longer targets (31). This finding appears to reflect greater processivity of the *Taq* polymerase (31). With *Taq*, targets of up to 500 base pairs in length can be amplified by using 3-min DNA synthesis reactions. By using 10-min DNA synthetic steps, the upper limit of fragment size lies between 2,000 and 2,500 base pairs (31).

Characteristics of PCR

Efficiency. The efficiency of PCR amplification depends on many variables. Some of the more important are (i) the number of cycles, (ii) the original amount of target, (iii) the length of the target sequence to be amplified, and (iv) the temperature of primer annealing and extension.

The first of these variables is the total number of cycles used. As noted above, the total amplification achieved is expressed as $(1 + x)^n$, where n is the number of cycles performed and x is the average per-cycle efficiency. However, as the concentration of products increases, the chance for product strands to reanneal to each other, rather than to primers, increases. Consequently, the later cycles of a PCR reaction have been observed to be less efficient than earlier cycles (2, 31). For example, beginning with 1 μg of genomic DNA and amplifying a segment of the β-globin gene, a 25-cycle reaction yielded 4.5×10^6-fold, a 30-cycle reaction yielded 8.9×10^6-fold, and a 35-cycle reaction yielded 1.7×10^7-fold amplification, with overall efficiencies of 85, 70, and 61%, respectively (31).

The second variable involves the amount of target to be amplified. Reactions beginning with relatively higher concentrations of target are less efficiently amplified. For example, a 10-cycle PCR reaction amplifying 10^{-15} mol of a 110-base-pair segment of the β-globin gene present in a plasmid template yielded a net amplification of 100-fold, or 62% overall efficiency (22). A 20-cycle reaction amplifying the same sequence, but using 1 μg of genomic DNA (containing only about 5×10^{-19} mol of β-globin target), yielded a 200,000-fold increase in copy number, which corresponds to about 85% overall efficiency (32).

The third variable is the length of the target sequence to be amplified. Efficiency is inversely related to the length of the target sequence amplified. Sequences of up to 2,000 base pairs have been amplified by using *Taq* DNA polymerase, but the efficiency of amplification of this size of region is considerably lower when compared with the amplification of sequences of regions of 100 to 300 base pairs (31). Thus, an effective amplification strategy is to select primers to amplify small but characteristics target regions (i.e., 100 to 300 base pairs).

Fourth, efficiency is also directly related to the temperature used for primer annealing and extension. Increasing the annealing temperature from 40 to 55°C has been observed to increase the overall sensitivity of β-globin amplification or detection by about 100-fold (31). This experiment involved dilution of normal genomic DNA with DNA from a cell line with a homozygous β-globin deletion. A 2-μg portion of total DNA was amplified in a 40-cycle PCR reaction. The DNA in the 10^{-6}-fold dilution was detectable when annealing was performed at 55°C, while only the DNA in the 10^{-4} dilution was detectable with 40°C annealing. In the 10^{-6} dilution experiment, the equivalent of only one β-globin sequence per 500,000 cells was present, highlighting the extraordinary sensitivity of detection obtained with PCR. The increased efficiency at elevated annealing temperatures may result

from improved reaction specificity (discussed below). Such high annealing and extension temperatures are possible only with the use of the thermostable *Taq* DNA polymerase.

Last, the concentration of the divalent cation Mg^{2+} is a critical factor influencing the efficiency of *Taq* polymerase-catalyzed PCR. While Mg^{2+} is required for *Taq* polymerase activity, excess Mg^{2+} can decrease the overall efficiency of amplification. The mechanism of this effect is unknown; the possibilities include an unusual Mg^{2+} dependency of the *Taq* DNA polymerase or a duplex-stabilizing effect of Mg^{2+} which interferes with target denaturation. In any event, the Mg^{2+} concentration during PCR should be empirically optimized for each target.

Specificity. The specificity of PCR depends primarily on the particular target sequence, the annealing temperature, the amount of DNA polymerase used, and the polymerization time of each cycle.

During the PCR process, nonspecific DNA synthesis will occur whenever a primer anneals to an incompletely homologous region of the original nucleic acid. However, unless the newly synthesized strand contains a region to which the second primer can anneal, it will not serve as a template in the following cycles. Thus, such "misprimed products" will accumulate in a linear, rather than exponential, fashion. An exception is mispriming within the portion of the target sequence bracketed by the two primers; such events will generate a partial-length product containing both primer binding sites and will amplify exponentially (22).

Elevated annealing temperatures will increase the stringency of hybridization between the oligonucleotide primers and the target nucleic acids. This reduces the amount of mispriming of nontarget sequences, conferring greater specificity on the amplification process. In the experiment described in the preceding section, which compared annealing at 40 versus 55°C, less nonspecific DNA was synthesized with 55°C annealing (31). Electrophoretic analysis of the amplification products permits verification of their size and homogeneity. This method of analysis also demonstrates that increasing amounts of enzyme and longer polymerization times result in relatively greater amounts of nonspecific DNA synthesis. These parameters (primer annealing temperature, amount of enzyme, and polymerization time of each cycle) should be empirically optimized for each target.

Finally, the specificity of PCR varies depending on the particular primers utilized and the composition of the background nucleic acid sequences. An experiment in which PCR-amplified sequences were cloned into an M13 vector demonstrates this point (33). β-Globin sequences were amplified and subsequently cloned into M13, a single-stranded DNA bacteriophage cloning vehicle. Although 80% of the clones contained sequences corresponding to the primers used for amplification, only 1% contained β-globin sequences. In contrast, when HLA-DQα sequences were similarly amplified and cloned, 20% of the clones contained the HLA-DQα sequences. The reason for this marked difference in specificity is unclear, but it has been hypothesized that these disparate results reflect different hybridization affinities of the primer pairs for nontarget background sequences in genomic DNA.

The specificity of PCR can be altered intentionally to enable amplification of closely related, but nonidentical sequences by the use of "degenerate" primers, oligonucleotide preparations that consist of a population of molecules with intentionally introduced sequence ambiguities (20). The short regions of sequence similarity within the reverse transcriptases of retroviruses and hepadnaviruses have been

exploited to design degenerate primers capable of efficiently amplifying reverse transcriptase sequences from both woodchuck hepatitis and duck hepatitis B viruses (10). Interestingly, these primers functioned much more efficiently if non-target-complementary 5' extensions encoding restriction endonuclease recognition sites were included in their sequences. Since the 5' extensions are incorporated into the amplification products during the initial PCR cycles (Fig. 2), subsequent cycles in which the products act as templates may benefit from a longer priming duplex, which confers improved annealing and extension efficiencies. However, this effect was not observed in a previous study similarly utilizing 5' extensions to facilitate the cloning of PCR fragments (33).

Error rates. It is important that amplification products represent faithful copies of the targeted sequences and thus have few misincorporated bases. This is especially true when the products are to be analyzed by nucleic acid hybridization or nucleotide sequence. The overall error frequency of a 30-cycle, *Taq* DNA polymerase-catalyzed PCR reaction used to amplify a DNA target has been estimated to be 0.25%, or one misincorporation per 400 bases. These data derive from an experiment in which 28 separate clones containing a 239-base-pair amplified product were sequenced; a total of 17 errors were found (31). The error frequency of amplifications of RNA targets has not yet been estimated but is expected to be higher, since the error frequency of reverse transcriptase (first step in the RNA amplification reaction) would be superimposed on the *Taq* DNA polymerase error rate.

GENERAL APPLICATIONS OF NUCLEIC ACID AMPLIFICATION IN VITRO

Nucleic Acid Sequence Detection Assays

Improvement of sensitivity and specificity of nucleic acid hybridization assays. As indicated previously, the use of nucleic acid sequence amplification in vitro prior to detection provides remarkable capabilities in sensitivity and specificity to nucleic acid hybridization assays. An illustrative example involves the use of allele-specific oligonucleotide probes for the detection of human β-globin alleles. Under appropriate hybridization conditions, these oligonucleotide probes are sensitive to single-base variations in the target sequence (5). Thus, such probes can be used to distinguish the A, S, and C allelic sequences of β-globin. Before the development of sequence amplification, such an analysis required 10 μg of genomic DNA (DNA from about 10^6 human cells) (5). In comparison, when the human β-globin target sequences are first amplified in vitro, using 25 cycles of PCR, only an amount equivalent to 33 ng of the original genomic DNA is required. The identification of the allele present in the genomic DNA sample is accomplished by using allele-specific oligonucleotide probes and a simple slot-blot hybridization protocol (27). The observed sensitivity of this amplification and hybridization protocol is approximately 100 times that of Southern analysis of unamplified DNA. A similar approach has been used for the rapid, prenatal diagnosis of sickle-cell anemia (8).

In addition to single-nucleotide changes, gross chromosomal alterations have been detected. Specifically, PCR has been used to detect a sequence specific to a chromosomal translocation event characteristic of follicular lymphomas (6, 18). Recently, as few as 1 in 10^6 cells carrying this translocation could be detected (6). PCR-based analysis detected

this sequence in a patient's remission marrow and blood, which appeared normal by Southern blot and morphologic analysis. Clearly, the detection of such correlative translocation events in patients undergoing therapy has significant implications for the effectiveness of treatment and ultimately for their prognosis.

Simplification of nucleic acid sequence detection. *Detection of specific sequences by direct visual inspection of gels.* Nucleic acid amplification in vitro allows detection of restriction fragment polymorphisms by direct visual inspection of electrophoretic gels, obviating the use of radioactive probes. The presence or absence of specific restriction endonuclease sites can be used as a marker for various alleles, such as mutations of hemophilia A in the factor VIII gene. This approach usually requires electrophoresis and Southern blotting of endonuclease-digested genomic DNA, followed by identification of the fragments with radioactive probes. When the portion of the factor VIII gene containing diagnostic restriction site polymorphisms for hemophilia A is amplified by a 30-cycle PCR reaction catalyzed by *Taq* polymerase, a discrete band is visible on ethidium-stained gels (16). Thus, the products of digestion with specific endonucleases can be analyzed directly by visual inspection.

Sequence detection without nucleic acid purification. PCR can be performed directly on crude cell lysates (16, 30). This further simplifies a detection assay by obviating DNA purification. It has been reported that a single-copy gene (β-globin) can be detected in as few as 75 cells by direct amplification of a cell lysate prepared simply by boiling (30). However, as the number of cells in the lysis preparation was increased, the intensity of the target-specific amplification signal first leveled off and then decreased, presumably because of inhibitors of the amplification reaction present in crude cell lysates. However, such inhibition is not always observed. In the study of the mutations of hemophilia A described in the preceding section, the PCR amplification reactions used alkaline lysates of chorionic villus cells (16). PCR amplification yielded discrete bands visible on electrophoretic gels. Thus, unlike the inhibition observed in the experiments involving the amplification of β-globin sequences, these results indicate that target amplification involving crude lysis preparation proceeded efficiently.

Nucleotide Sequencing of Targeted Sequences

PCR has been utilized for direct sequence analysis of genomic DNA. This method is useful when the objective is to analyze a limited region of DNA for sequence variations. The method involves PCR amplification of a region of the target sequence, followed by annealing with a third, radiolabeled primer to the isolated, amplified DNA. This primer is extended in the presence of dideoxynucleotides for sequencing by the chain termination method. This method permits the rapid analysis of allelic sequence variations among individuals without requiring the laborious process of cloning each individual sequence. This approach has been used to characterize the sequences of alleles of β-thalassemia (37), the c-*ras* oncogene (21), and human mitochondrial DNA (38).

An enhanced PCR protocol has been developed to facilitate the direct sequencing of the amplified products (35). By using one PCR primer having the promoter sequence for the bacteriophage T7 RNA polymerase on its 5' portion, the amplified products which contain T7 promoter sequences are transcribed after multiple cycles of PCR amplification. The resulting single-stranded RNA is then annealed to a primer

for sequencing. The transcription step enhances the overall amplification and generates single-stranded template for the sequencing reaction.

Cloning of Targeted Sequences

PCR can be used to enhance the efficiency of cloning a specific sequence of interest (33). Restriction endonuclease-specific sequences can be included on the 5' portions of the PCR primers and are thus built into the ends of the amplified sequence. The limitations of this method are the length of the sequence that can be amplified (currently about 2,000 base pairs) and the need to know in advance the sequences flanking the segment to be cloned. If the object is to clone a known sequence, however, this method obviates the required construction and screening of full genomic libraries.

APPLICATION OF SEQUENCE AMPLIFICATION TO THE DETECTION AND STUDY OF HUMAN PATHOGENIC RETROVIRUSES

Application to HIV-1

The problem of detection of HIV-1. Detection of HIV-1 requires a highly sensitive assay because of the exceedingly small quantity of virus and viral products present in clinical samples. The fraction of cells expressing HIV-1 is estimated to be $<100/10^6$, and in some samples $<10/10^6$, both for peripheral blood mononuclear cells (PBMCs) and biopsy samples of lymph nodes. Furthermore, the level of viral expression in these cells is quite low, perhaps 20 to 300 RNA copies per cell, as determined by in situ hybridization (14). The presence of viral DNA in clinical samples has been assessed by conventional methods, but the sensitivity is poor. Southern blots of endonuclease-digested DNA probed with nick-translated, proviral fragments yielded positive results in only 8 of 56 samples (14%) of PBMCs and lymph nodes from patients with acquired immunodeficiency syndrome (AIDS) and AIDS-related complex (34). Direct detection of HIV-1 RNA from PBMCs of infected individuals yields improved sensitivity, approaching that of coculture assays (24, 26). However, signal intensity is low. Thus, a more sensitive assay is needed to study the expression of HIV during the natural history of infection and during antiviral therapy.

Detection of HIV-1 in clinical specimens by nucleic acid amplification. *DNA detection.* Recently, the use of sequence amplification for the detection of HIV-1-specific DNA has been reported (17, 19, 23). The sensitivity of this method far exceeds that of conventional methods. In a recent study, use of the PCR amplification protocol permitted the detection of HIV sequences in 100% of the culture-positive patients and in 64% (7 of 11) of the seropositive but culture-negative patients (23). No false-positive results occurred. For each of these samples, 1 μg of DNA from PBMCs was amplified by using 35 cycles of PCR, and the products were analyzed by oligomer restriction. In this technique of oligomer restriction, a radiolabeled oligonucleotide probe is annealed to the amplification products in solution and then cut with a specific endonuclease to generate a diagnostic fragment, which is resolved by electrophoresis (29).

Interestingly, this study demonstrates the potential of the PCR protocol to generate false-negative results, presumably caused by variations in the HIV-1 sequence (23). Four different primer pairs designed to amplify four distinct segments of the HIV-1 genome were tested. The sensitivity of

any one primer pair in detecting HIV-1 sequences in culture-positive patients was close or equal to 100%. However, for seropositive but culture-negative patients, different primer pairs had a range of sensitivities, from 2 of 11 to 5 of 11 positive results. To achieve maximal sensitivity in this seropositive but culture-negative group, a given sample required amplification with multiple primer pairs. Two primer pairs from the *env* gene were sufficient to encompass the entire range of positive results. HIV-1 sequence variation may account for the inability of a particular primer or probe set to generate a positive signal; such sequence variation could reduce the efficiency of primer annealing for amplification or the efficiency of probe hybridization. Sequence heterogeneity is a well-documented property of the HIV-1 genome (13, 36). Thus, for HIV-1 confirmation assays, it would be necessary to pick primer and probe sequences that are well conserved among HIV-1 isolates. Even so, it appears that, for HIV-1, multiple regions of the genome must be amplified to obtain maximal sensitivity.

It should be noted that the detection method used in the study cited is especially sensitive to target sequence variations. As noted above, oligomer restriction depends on the action of a restriction endonuclease. A single-base change in the amplified target (introduced by strain variation or by errors in amplification) at the sequence recognized by this enzyme can cause a false-negative result due to failure to cut the target or probe duplex. This is the principal reason why the oligomer restriction detection method will likely be replaced by an oligomer hybridization method, which requires only hybridization of the ^{32}P-labeled probe.

The extraordinary sensitivity of this amplification-based assay for HIV-1 can lead to difficulties in the interpretation of the data. For example, the meaning of the positive results from sequence amplification in a seropositive but viral culture-negative individual is open to speculation (23). It appears to indicate a sensitivity greater than that of coculture assays. However, the state of activity of the viral nucleic acids in these patients is not clear. For example, positive results could be generated by the presence of partial, defective proviral sequences which are incapable of supporting viral replication.

Similar difficulties were encountered in interpreting the results of a PCR-based assay applied to four healthy, HIV-1 antibody-positive homosexual men who became seronegative during follow-up (9). All of the individuals were positive by the PCR assay at least once during three follow-up determinations, made at 6-month intervals following loss of demonstrable HIV-1 antibody. Two individuals became negative on the last two determinations, indicating that they may have cleared their viral load below the limits of detection by PCR. The significance of the persistent viral sequences in the other two patients will be clarified only after continued longitudinal follow-up.

Clearly, early HIV-1 detection can be facilitated by amplification-based assays. Two patient groups are of major clinical importance in this regard: (i) individuals in the window between exposure and the development of antibody; and (ii) infants of seropositive mothers in whom the presence of HIV-1 cannot be determined by serological analysis during the first 6 months of life because of the presence of maternal antibodies in their sera.

A recent study addressed the first patient group by analyzing the DNA from PBMCs of seronegative individuals who were sexual partners of seropositive individuals (19). Of 16 such individuals, 5 were positive by PCR amplification of a *gag* gene sequence, followed by a format that included

slot-blot, ^{32}P-labeled oligonucleotide, and hybridization detection. The results were confirmed by gel electrophoresis of the amplification products and detection of the *gag*-specific sized fragment. These individuals were negative for HIV-1 antibody by enzyme-linked immunosorbent assay and Western blot (immunoblot), as well as being negative for p24 antigen. All five individuals were reportedly positive by a PCR assay repeated 2 to 3 months later, although all remained negative by antibody and antigen assays. The significance of these findings can be evaluated only after continued longitudinal study has defined the natural history of these individuals.

RNA detection. An assay for HIV-1 RNA has an advantage over DNA detection in that the presence of viral RNA indicates ongoing viral replication. The presence of DNA sequences, on the other hand, gives no information regarding the state of viral replicative activity. Amplification of RNA sequences can be performed by target-specific complementary DNA synthesis, followed by PCR, as described above. This approach has been applied to PBMCs of seropositive patients (15). Cells were lysed in guanidinium isothiocyanate (a chaotropic salt solution), and the DNA and RNA were separated by centrifugation on a cesium chloride gradient. For each sample, DNA and RNA fractions were separately processed for amplification of a *gag* gene sequence. Detection of this sequence was accomplished by oligomer restriction. Of 18 patients, 17 were positive for *gag* sequence DNA. These patients were seropositive and included individuals who had AIDS-related complex and AIDS, as well as no symptoms. Of the 17 DNA-positive patients, 11 were positive for RNA. These RNA-positive patients included several who were receiving various antiviral agents, including azidothymidine, acyclovir, or Ampligen (a mismatched, double-stranded RNA).

Six patients were negative for RNA but positive for DNA. Assuming the absence of experimental error, this result supports the concept of transcriptional latency in the cells assayed; that is, proviral DNA is present, but no detectable viral message is being synthesized. It is curious, therefore, that five of six of these individuals were symptomatic, four with AIDS-related complex and one with AIDS. Only one was asymptomatic, and none were receiving antiviral therapy. It is a paradox that these individuals, who were clinically ill with HIV infection, had no detectable RNA expression, yet viral DNA was demonstrable. Two explanations are apparent, both of which have important implications for the design of future studies. One is that HIV expression in vivo is subject to bursts of viral replication, interspersed with periods of transcriptional inactivity. If this is true, then the PCR analysis must include multiple-time-point data for each individual, to follow a fluctuating viral load over a sufficient period of time to establish an overall trend. The second possibility is that the mononuclear cells in the peripheral blood are not representative of the overall state of HIV activity in vivo. Further analysis of RNA and DNA signals over time in patients whose clinical course and serological status are well characterized will be needed to clarify these issues.

Applications of nucleic acid amplification in vitro in the study of HIV-1 pathogenesis. Besides the early detection of HIV-1 nucleic acid sequences in clinical samples, several promising uses of in vitro nucleic acid amplification for the study of HIV-1 pathogenesis are possible. These approaches may also be used for the study of any viral pathogen.

Quantitation of HIV-1 nucleic acids. The development of a quantitative assay for HIV-1-specific sequences based on
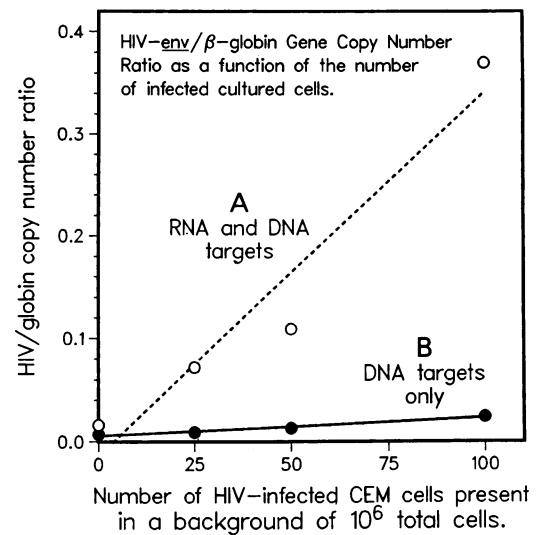

FIG. 3. HIV *env*/β-globin gene copy number as a function of number of infected cultured cells. The value of the HIV-1 *env*/β-globin gene ratio is linear over 0 to 100 infected CEM cells (a CD4-positive human T-cell line) present in a background of 10^6 cells. Total nucleic acids were isolated from 10^6 cells of each dilution. These samples were divided in half. Sample A was amplified with an initial reverse transcription reaction, followed by PCR, amplifying both DNA and RNA targets. Sample B was amplified without reverse transcription; only DNA targets were amplified. HIV *env* and β-globin sequences were amplified simultaneously in each sample. After amplification, aliquots of each sample were slot-blotted and hybridized separately for HIV *env* and β-globin sequences, using ^{32}P-labeled oligonucleotide probes. Individual slots were counted in a scintillation counter. Values were corrected for the relative specific activities of HIV and β-globin probes, and a relative ratio of HIV-1 *env* copy number to single-copy gene β-globin was calculated.

sequence amplification appears achievable. This is possible under conditions in which nucleic acid isolation is reproducible, the efficiency of amplification is relatively constant, and the detection method can measure different yields of amplified sequence. Such an assay would be valuable in the study of the natural history of HIV-1 infection and in the evaluation of the effects of antiviral therapy. To some degree, variations among samples can be normalized by the use of an internal nucleic acid standard. This quantitative assay involves the coamplification of an endogenous nucleic acid sequence whose copy number per cell is invariant. For example, amplification of a single-copy gene (β-globin) can be performed simultaneously with HIV-1-specific sequences. The signal for β-globin can then be used to standardize the HIV-1 signal intensities among samples, which may differ in the amount of nucleic acid isolated or in amplification efficiencies. Preliminary results with this approach indicate a linear relationship between the intensity of the HIV-1 amplification signal and the number of HIV-1-infected cultured cells assayed (Fig. 3). A similar approach has been used to quantitate the amount of dystrophin transcript in various tissues as a relative value compared with the amount of aldolase A transcript (2).

For reproducible quantitation, the target sequence to be measured and the internal standard with which it is compared must be amplified with similar efficiencies. Since the efficiency of the PCR reaction decreases with the number of cycles performed, it is important to document that both the target and internal standard sequences are in the exponential
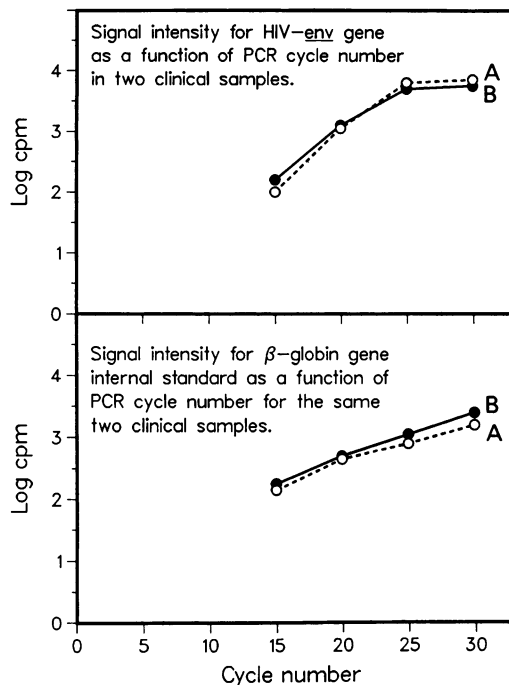
FIG. 4. Signal intensity for HIV *env* gene (top) or β-globin gene (bottom) as a function of PCR cycle number. The increase in signal intensity of HIV *env* sequences in two samples of nucleic acids from the PBMCs of an AIDS patient is exponential until the 25th cycle of PCR, after which it plateaus. β-Globin sequences are exponential over 15 to 30 cycles. Clearly, the HIV *env*/β-globin ratio will be artifactually decreased if the detection assay is performed after 30 cycles rather than 25. Total nucleic acids were isolated from 10^6 PBMCs isolated from an AIDS patient at two different times (A and B). Amplification of both RNA and DNA sequences was performed by a reverse transcription reaction, followed by PCR. Aliquots of amplification products taken at 15, 20, 25, and 30 cycles were slot-blotted and hybridized separately for HIV *env* and β-globin. (A 20-fold-greater amount of material was assayed for HIV compared with β-globin.) Individual slots were counted in a scintillation counter, and the log counts per minute were plotted against cycle number.

range of amplification during the assay (2). An example of different amplification efficiencies is shown in Fig. 4, in which the log of the amount of amplified HIV-1 *env* and β-globin sequences is plotted against the cycle number of the PCR reaction. The two clinical samples represent the total nucleic acids isolated from 500,000 PBMCs of an AIDS patient at different times. The signal for HIV-1 *env* sequence plateaus between 25 and 30 cycles of PCR. The signal for β-globin appears linear (exponential) over this range. For the ratio of HIV-1/β-globin after amplification to reflect the original copy number ratio in the isolated nucleic acids, the detection assay should be performed after 25 cycles of PCR, when the amplification of both target (HIV-1) and standard (β-globin) is in the exponential range.

Assay of transcriptional activity. An assay that assesses viral transcriptional activity can be designed by using PCR amplification. Specific sequences in RNA molecules can be amplified as described above. If reverse transcriptase is used in the first polymerization step of a PCR reaction, HIV-1 sequences present as both RNA and DNA will be amplified (Fig. 3). If reverse transcriptase is omitted, only DNA sequences will be amplified. In principle, comparison of parallel samples amplified with and without reverse tran-

scription yields an index of viral transcriptional activity. For an RNA virus such as HIV-1, transcriptional activity is a direct reflection of viral replication. Various factors, including T-cell activation and transactivation by DNA viruses, are known to stimulate HIV-1 replication in vitro (11, 39). The question of whether such viral activation occurs during the natural history of HIV-1 infection may be approached by using an in vitro nucleic acid amplification-based assay.

Viral messenger RNA splicing analysis. PCR can be used to detect and quantitate specific species of viral RNA that are generated by differential RNA splicing. The result of RNA splicing is the juxtaposition of nucleic acid sequences that are several thousand base pairs apart in proviral DNA or genomic RNA. If primers are positioned on either side of a particular splice junction, the sequence of the spliced RNA is specifically amplified. Although the primers will anneal to the same sequences in proviral DNA and in RNA molecules from which the intervening intron has not been spliced out, effective amplification with these molecules as templates will not occur because of the great distance between the two primers. Amplification of a HIV-1 messenger RNA sequence appropriately spliced for *tat* expression has been reported (15). Furthermore, different splice events occurring between the primer sequences should yield amplification products of characteristic length. These fragments can be isolated and sequenced as described above. This approach can be used to identify and characterize novel splice sites.

HIV-1 has a single transcription start site, but has three major species of RNA generated from the primary transcript by RNA slicing (10). The viral gene *rev* (*trs/art*) has been implicated as a regulator of splicing by an as yet uncharacterized mechanism. In the absence of *rev*, spliced species encoding viral structural proteins and full-length genomic RNA are not produced, although viral regulatory genes are expressed (10). By using sequence amplification, it may be possible to ask whether such a switch in RNA processing by differential splicing actually occurs during the natural process of viral replication.

HIV sequencing. Direct sequence analysis of amplified products can be used to speed the acquisition of data regarding sequence variation among HIV-1 isolates. HIV-1 isolates prepared sequentially from specific individuals as their disease progressed have shown increasing virulence in vitro (3). Direct sequencing of amplified portions of these isolates could facilitate the process of sequence comparison, helping to define which, if any, sequences correspond to increased virulence.

It has become clear that HIV-1 isolated from an individual at any point in time is composed of a collection of genomes with significant nucleotide sequence variation (28). With direct nucleotide sequence analysis, PCR fragments amplified from nucleic acids isolated from an infected individual are expected to display sequence ambiguities at the positions that are variable in the original viral genomes. This result might be useful as an approach to ascertain conserved and variable regions of the viral genome. If a specific, unambiguous sequence is sought, however, it would be necessary to determine the nucleotide sequence of individual clones derived from the amplified region.

Application to HTLV-I

Sequence amplification by PCR has recently been applied to the detection of human T-cell lymphoma virus type I (HTLV-I) in clinical specimens. This human retrovirus is associated with both adult T-cell leukemia or lymphoma and

a neurologic syndrome with several names, including chronic progressive myelopathy and tropical spastic paraparesis. Like HIV-1, this retrovirus is pathogenic at low copy numbers. Therefore, direct detection of viral components in clinical samples is difficult. In a study of patients with chronic progressive myelopathy, a PCR-based detection system proved vastly more sensitive than conventional Southern blot assays, but no more sensitive than detection of antibodies to HTLV-I by enzyme-linked immunosorbent assay or Western blot analysis (1). Of 11 antibody-positive patients with chronic progressive myelopathy, none had detectable HTLV-I sequences in peripheral blood lymphocytes as determined by Southern hybridization analysis. However, all of these patients were positive by a PCR-based assay. Five antibody-negative patients were all negative by PCR as well.

The PCR assay was able to confirm the presence of HTLV-I-specific sequences in a biopsy specimen from a patient previously believed to have Hodgkin's lymphoma (7). This patient was negative for HTLV-I antibodies by enzyme-linked immunosorbent assay and Western blot. Biopsy specimens were negative by Southern analysis, yet positive by PCR. Notably, biopsy samples were also positive by immunohistochemical staining with a monoclonal antibody to the HTLV-I p19 *gag* antigen.

CLINICAL DIAGNOSIS AND SEQUENCE AMPLIFICATION

PCR has been proposed as a method to identify infected patients or blood donors before a detectable seroconversion or to adjudicate between conflicting serological assays when the question of infection in an individual is ambiguous. This is an inappropriate and premature application of a promising technique which must itself first be validated in a systematic fashion with well-characterized clinical specimens in blinded studies. An additional potential application of the method is the discrimination between HTLV-I and HTLV-II infection or between HIV-1 and HIV-2 infection. In certain individuals, antibody assays may be unable to discriminate which virus type is responsible for a confirmed, positive test; however, type-specific nucleic acid sequences can be identified for use in PCR assays.

It is premature to consider the methodology available in 1989 for clinical use. Techniques and reagents are in rapid evolution. Many practical hurdles, such as false-positive results due to contamination of reagents and equipment with sequences, have to be addressed. Yet sequence amplification may facilitate the adaptation of nucleic acid hybridization assays to routine use in the clinical laboratory. To date, widespread use of nucleic acid diagnostics has not occurred, primarily for three reasons (25, 27). One is the need for nucleic acid isolation. The second is the cumbersome method of immobilizing nucleic acids onto solid supports, followed by probe hybridization and washing. Last, the use of radioactive probes is undesirable for both safety and cost reasons. Sequence amplification can address these issues. As described above, PCR can be performed directly on crude cell lysates, obviating nucleic acid purification. New methods of nucleic acid immobilization, or "capture," are on the horizon (12), as are nonisotopic probes. Any relative loss in sensitivity due to either of these approaches can be compensated for by sequence amplification prior to detection. Thus, the future role of sequence amplification in clinical viral diagnostics promises to be important.

CONCLUSION

The advent of a technique for target-specific nucleic acid sequence amplification in vitro has had and will continue to have a dramatic impact on both the study of viral pathogenesis and the routine detection of viral pathogens in the clinical laboratory. Although many of the specific details of the methodology discussed in this review will doubtless be rendered obsolete, the general advantages of this approach will remain. First, nucleic acid amplification in vitro prior to detection will dramatically improve the sensitivity and specificity of any nucleic acid detection assay. For the study of viral pathogenesis, this will enable important questions to be answered for which the test sensitivity has previously been insufficient. For the clinical laboratory, the improved sensitivity and specificity conferred by in vitro nucleic acid amplification may facilitate the development of new detection techniques which may be more suitable for routine use than current conventional methods. Second, the ability to replicate a nucleic acid sequence in vitro, in a sense, "cell-free molecular cloning" (31), will greatly speed the manipulations required in the investigation of the molecular biology of viral pathogens.

The development of nucleic acid probe diagnostics as facilitated by nucleic acid amplification in vitro is in an early stage. Many problems must be faced before a standard assay for pathogens such as HIV-1 is achieved. Most prominent among these difficulties are the standardization of sample preparation, amplification (for example, the choice of the genomic region to be amplified), and, finally, the detection format itself. The potential rewards of this approach are motivating intense, ongoing investigations.

LITERATURE CITED

1. **Bhagavati, S., G. Ehrlich, R. W. Kula, S. Kwok, J. Sninsky, V. Udani, and B. J. Poiesz.** 1988. Detection of human T-cell lymphoma/leukemia virus type I DNA and antigen in spinal fluid and blood of patients with chronic progressive myelopathy. N. Engl. J. Med. 318:1141–1147.
2. **Chelly, J., J.-C. Kaplan, P. Maire, S. Gautron, and A. Kahn.** 1988. Transcription of the dystrophin gene in human muscle and non-muscle tissues. Nature (London) 333:858–860.
3. **Cheng-Mayer, C., D. Seto, M. Tateno, and J. A. Levy.** 1988. Biological features of HIV-1 that correlate with virulence in the host. Science 240:80–82.
4. **Chien, A., D. B. Edgar, and J. M. Trela.** 1976. Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus.* J. Bacteriol. 127:1550–1557.
5. **Conner, B. J., A. A. Reyes, C. Morin, K. Itakura, R. L. Teplitz, and R. B. Wallace.** 1983. Detection of sickle cell β^s-globin allele by hybridization with synthetic oligonucleotides. Proc. Natl. Acad. Sci. USA 80:278–282.
6. **Crescenzi, M., M. Seto, G. P. Herzig, P. D. Weiss, R. C. Griffith, and S. J. Korsmeyer.** 1988. Thermostable DNA polymerase chain amplification of t(14;18) chromosome breakpoints and detection of minimal residual disease. Proc. Natl. Acad. Sci. USA 85:4869–4873.
7. **Duggan, D. B., G. D. Ehrlich, F. P. Davey, S. Kwok, J. Sninsky, J. Goldberg, L. Baltrucki, and B. J. Poiesz.** 1988. HTLV-I-induced lymphoma mimicking Hodgkin's disease. Diagnosis by polymerase chain reaction amplification of specific HTLV-I sequences in tumor DNA. Blood 71:1027–1032.

8. Embury, S. H., S. J. Scharf, R. K. Saiki, M. A. Gholson, M. Golbus, N. Arnheim, and H. A. Erlich. 1987. Rapid prenatal diagnosis of sickle cell anemia by a new method of DNA analysis. N. Engl. J. Med. 316:656–661.

9. Farzadegan, H., M. A. Polis, S. M. Wolinsky, C. R. Rinaldo, Jr., J. J. Sninsky, S. Kwok, R. L. Griffith, R. A. Kaslow, J. P. Phair, B. F. Polk, and A. J. Saah. 1988. Loss of human immunodeficiency virus type 1 (HIV-1) antibodies with evidence of viral infection in asymptomatic homosexual men. Ann. Intern. Med. 108:785–790.

10. Feinberg, M. B., R. F. Jarrett, A. Aldovini, R. C. Gallo, and F. Wong-Staal. 1986. HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. Cell 46:807–817.

11. Gendelman, H. E., W. Phelps, L. Feigenbaum, J. M. Ostrove, A. Adachi, P. M. Howley, G. Khoury, H. S. Ginsberg, and M. A. Martin. 1986. Trans-activation of the human immunodeficiency virus long terminal repeat sequence by DNA viruses. Proc. Natl. Acad. Sci. USA 83:9759–9763.

12. Gingeras, T. R., D. Y. Kwoh, and G. R. Davis. 1987. Hybridization properties of immobilized nucleic acids. Nucleic Acids Res. 15:5373–5390.

13. Hahn, B. H., M. A. Gonda, G. M. Shaw, M. Popovic, J. A. Hoxie, R. C. Gallo, and F. Wong-Staal. 1985. Genomic diversity of the acquired immune deficiency syndrome virus HTLV-III: different viruses exhibit greatest divergence in their envelope genes. Proc. Natl. Acad. Sci. USA 82:4813–4817.

14. Harper, M. E., L. M. Marselle, R. C. Gallo, and F. Wong-Staal. 1986. Detection of lymphocytes expressing human T-lymphotropic virus type III in lymph nodes and peripheral blood from infected individuals by in situ hybridization. Proc. Natl. Acad. Sci. USA 83:772–776.

15. Hart, C., T. Spira, J. Moore, J. Sninsky, G. Schochetman, A. Lifson, J. Galphin, and C.-Y. Ou. 1988. Direct detection of HIV RNA expression in seropositive subjects. Lancet ii:596–599.

16. Kogan, S. C., M. Doherty, and J. Gitschier. 1987. An improved method for prenatal diagnosis of genetic diseases by analysis of amplified DNA sequences. Application to hemophilia A. N. Engl. J. Med. 317:985–990.

17. Kwok, S., D. H. Mack, K. B. Mullis, B. Poiesz, G. Ehrlich, D. Blair, A. Friedman-Kien, and J. J. Sninsky. 1987. Identification of human immunodeficiency virus sequences by using in vitro enzymatic amplification and oligomer cleavage detection. J. Virol. 61:1690–1694.

18. Lee, M.-S., K.-S. Chang, F. Cabanillas, E. J. Freireich, J. M. Trujillo, and S. A. Stass. 1987. Detection of minimal residual cells carrying the t(14;18) by DNA sequence amplification. Science 237:175–178.

19. Loche, M., and B. Mach. 1988. Identification of HIV-infected seronegative individuals by a direct diagnostic test based on hybridisation to amplified viral DNA. Lancet ii:418–421.

20. Mack, D. H., and J. J. Sninsky. 1988. A sensitive method for the identification of uncharacterized viruses related to known virus groups: hepadnavirus model system. Proc. Natl. Acad. Sci. USA 85:6977–6981.

21. McMahon, G., E. Davis, and G. N. Wogan. 1987. Characterization of c-Ki-ras oncogene alleles by direct sequencing of enzymatically amplified DNA from carcinogen-induced tumors. Proc. Natl. Acad. Sci. USA 84:4974–4978.

22. Mullis, K. B., and F. A. Faloona. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol. 155:335–350.

23. Ou, C.-Y., S. Kwok, S. W. Mitchell, D. H. Mack, J. J. Sninsky, J. W. Krebs, P. Feorino, D. Warfield, and G. Schochetman. 1988. DNA amplification for direct detection of HIV-1 in DNA of peripheral blood mononuclear cells. Science 239:295–297.

24. Pellegrino, M. G., M. Lewin, W. A. Meyer III, R. S. Lanciotti, L. Bhaduri-Hauck, D. J. Volsky, K. Sakai, T. M. Folks, and D. Gillespie. 1987. A sensitive solution hybridization technique for detecting RNA in cells: application to HIV in blood cells. BioTechniques 5:452–459.

25. Richman, D. D., P. H. Cleveland, D. C. Redfield, M. N. Oxman, and G. M. Wahl. 1984. Rapid viral diagnosis. J. Infect. Dis. 149:298–310.

26. Richman, D. D., J. A. McCutchan, and S. A. Spector. 1987. Detecting human immunodeficiency virus RNA in peripheral blood mononuclear cells by nucleic acid hybridization. J. Infect. Dis. 156:823–827.

27. Richman, D. D., and G. M. Wahl. 1986. Nucleic acid probes to detect viral diseases, p. 301–309. In A. N. Notkins and M. B. A. Oldstone (ed.), Concepts in viral pathogenesis, vol. 2. Springer-Verlag, New York.

28. Saag, M. S., B. H. Hahn, J. Gibbons, Y. Li, E. S. Parks, W. P. Parks, and G. M. Shaw. 1988. Extensive variation of human immunodeficiency virus type-1 in vivo. Nature (London) 334:440–447.

29. Saiki, R. K., N. Arnheim, and H. A. Erlich. 1985. A novel method for the detection of polymorphic restriction sites by cleavage of oligonucleotide probes: application to sickle-cell anemia. Bio/Technology 3:1008–1012.

30. Saiki, R. K., T. L. Bugawan, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1986. Analysis of enzymatically amplified β-globin and HLA-DQα DNA with allele-specific oligonucleotide probes. Nature (London) 324:163–166.

31. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491.

32. Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230:1350–1354.

33. Scharf, S. J., G. T. Horn, and H. A. Erlich. 1986. Direct cloning and sequence analysis of enzymatically amplified genomic sequences. Science 233:1076–1078.

34. Shaw, G. M., B. H. Hahn, S. K. Arya, J. E. Groopman, R. C. Gallo, and F. Wong-Staal. 1984. Molecular characterization of human T-cell leukemia (lymphotropic) virus type III in the acquired immune deficiency syndrome. Science 226:1165–1171.

35. Stoflet, E. S., D. D. Koeberl, G. Sarkar, and S. S. Sommer. 1988. Genomic amplification with transcript sequencing. Science 239:491–494.

36. Willey, R. L., R. A. Rutledge, S. Dias, T. Folks, T. Theodore, C. E. Buckler, and M. A. Martin. 1986. Identification of conserved and divergent domains within the envelope gene of the acquired immunodeficiency syndrome retrovirus. Proc. Natl. Acad. Sci. USA 83:5038–5042.

37. Wong, C., C. E. Dowling, R. K. Saiki, R. G. Higuchi, H. A. Erlich, and H. H. Kazazian, Jr. 1987. Characterization of β-thalassaemia mutations using direct genomic sequencing of amplified single copy DNA. Nature (London) 330:384–386.

38. Wrischnik, L. A., R. G. Higuchi, M. Stoneking, H. A. Erlich, N. Arnheim, and A. C. Wilson. 1987. Length mutations in human mitochondrial DNA: direct sequencing of enzymatically amplified DNA. Nucleic Acids Res. 15:529–541.

39. Zagury, D., J. Bernard, R. Leonard, R. Cheynier, M. Feldman, P. S. Sarin, and R. C. Gallo. 1986. Long-term cultures of HTLV-III infected cells: a model of cytopathology and T-cell depletion in AIDS. Science 231:850–853.