Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 6977–6981, September 1988
Microbiology

A sensitive method for the identification of uncharacterized viruses related to known virus groups: Hepadnavirus model system

(polymerase chain reaction/reverse transcriptase/retroviruses)

DAVID H. MACK AND JOHN J. SNINSKY*

Cetus Corporation, Department of Infectious Diseases, 1400 Fifty-Third Street, Emeryville, CA 94608

Communicated by Joshua Lederberg, April 14, 1988 (received for review December 21, 1987)

ABSTRACT **Amino acid sequence similarity of the reverse transcriptases encoded by retroviruses and hepadnaviruses was first reported by Toh, H., Hayashida, H. & Miyata, T. (1983) *Nature (London)* 305, 827–829. The regions of similarity extend over a small number of amino acids and require the introduction of gaps through the open reading frame. By using an octapeptide region as the sole criterion for "taxonomic" classification, we have grouped the oncoviruses into two distinct categories and the lentiviruses and hepadnaviruses into two additional groupings. This classification suggests that murine and feline leukemia viruses may be more closely related to the viruses that are associated with leukemia in primates and cattle than had been appreciated. We have exploited a portion of this region because of the minimal translational codon degeneracy of the conserved residues. Unique oligonucleotides from this region have been designed and used in the primer-directed *in vitro* DNA amplification of the hepadnaviruses as a model system. In addition, mixtures of oligonucleotides with various sequences but of the same length were demonstrated to be efficient primers. The amplification procedure enabled dramatic increases in sensitivity and coincident detection of mammalian and avian genomes. This approach will be a valuable tool to detect and characterize members of viral groups. In addition, since short stretches of similarity have been frequently identified in related but distinct genes, such an approach could prove a valuable asset to molecular studies in general.**

Many investigators have speculated that to date only a fraction of the viruses responsible for human disease have been identified. The association of disease with the human lymphotropic viruses [human T-cell leukemia viruses types I (HTLV-I, ref. 1) and II (HTLV-II, ref. 2) and human immunodeficiency virus types 1 (HIV-1, refs. 3 and 4) and 2 (HIV-2, ref. 5)] as well as hepatitis B virus (HBV, refs. 6 and 7) suggests that viruses that replicate through an RNA intermediate represent a particularly important class of human pathogens. The identification of viruses generally requires either the presence of a substantial number of discernible particles in the infected organism or the *in vitro* propagation of the virus. The detection of viruses with host specificities restricted to cell types not readily cultured has proven particularly difficult. Nucleic acid hybridization and immunological cross-reaction have been used to detect viruses that are related to characterized viruses. Efforts to identify viral nucleic acid from potentially infected eukaryotic cells have proven problematic, due to the small fraction of infected cells and the minimal complementarity between target and probe sequences. Although notable complementarity has subsequently been identified for several viruses, the regions of similarity frequently extend over a small number of nucleotides. The complexity of eukaryotic DNA poses additional hurdles since the use of hybridization at decreased stringency typically results in unacceptable backgrounds.

We reasoned that selective amplification of invariant or highly conserved regions of viral nucleic acid would decrease the complexity of the DNA to be probed and allow detection of viral genomes related to known viruses. We identify here a region of the reverse transcriptase gene of retroviruses and hepadnaviruses that can serve as a template for the oligonucleotide-primed amplification procedure that may permit detection of uncharacterized viral nucleic acid sequences.

MATERIALS AND METHODS

Enzymatic Amplification. To obtain amplification of specific DNA regions the enzymatic process known as the polymerase chain reaction (PCR) was applied (8, 9). For PCR targets, recombinant plasmids containing the entire woodchuck hepatitis virus (WHV) (10) and duck hepatitis B virus (DHBV) (11) genomes were added to human genomic DNA. Reaction mixtures (100 μl) containing 1 μg of genomic DNA (with and without addition of plasmid DNA at one copy per genome) were prepared as described (9).

Analysis of Amplified DNA. One fifteenth of the reaction mixture for the PCR was used for ethidium bromide/gel electrophoresis and Southern analysis as described (9). The hybridization solution for the oligonucleotide probe, designated MD09, contained 3× SSPE (1× SSPE = 0.18 M NaCl/10 mM sodium phosphate, pH 7.4/1 mM EDTA), 5× Denhardt's solution (1× Denhardt's solution = 0.02% polyvinylpyrrolidone/0.02% Ficoll/0.02% bovine serum albumin), 30% (vol/vol) formamide, and 0.5% $NaDodSO_4$. The blot was incubated with 0.5 pmol of labeled probe per ml at 42°C for 3–5 hr. Excess probe was removed by one wash with 2× SSPE/0.1% $NaDodSO_4$ and by an additional wash with 0.2× SSPE/0.1% $NaDodSO_4$ 10 min at room temperature. The filter was then blotted dry and autoradiographed overnight at −70°C with a single intensifying screen (DuPont Lightning Plus).

Cloning of Amplified Product. After analysis of the amplification reaction, one-half of the total reaction volume was digested with 40 units of *Bam*HI and 40 units of *Hin*dIII as suggested by the vendor (New England Biolabs). Digested DNA was purified and concentrated before ligation as described (12) by using the Geneclean system (BIO 101). All subsequent steps were performed as described (13).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HTLV-I and -II, human T-cell lymphotropic viruses types I and II, respectively; HIV-1 and -2, human immunodeficiency virus types 1 and 2, respectively; PCR, polymerase chain reaction; WHV, woodchuck hepatitis virus; DHBV, duck hepatitis B virus; BLV, bovine leukemia virus; MuLV, murine leukemia virus; FeLV, feline leukemia virus; HBV, hepatitis B virus; GSHV, ground squirrel hepatitis virus.

*To whom reprint requests should be addressed.

RESULTS

Strategy. Frequently, regional nucleic acid sequence similarity among viruses is evident only after comparative, computer-assisted analyses. The conserved regions typically encode portions of functionally important proteins in the viral life cycle.

Toh *et al.* (14) initially noted that the amino acid sequences of the reverse transcriptases of Rous sarcoma virus and murine leukemia virus (MuLV) and the polypeptide predicted by the largest open reading frame of HBV displayed minimal but notable similarity. Other laboratories (15–17) have extended this observation to other members of the oncoviruses and lentiviruses. A conserved tetrapeptide, Tyr-Met-Asp-Asp, of these polymerases has been noted. Since the sequence of this region is invariant among all exogenous viruses except MuLV and feline leukemia virus (FeLV) in which a valine substitutes for methionine, we have designated it the "generic" primer binding site. The report (18) that polypeptides with mutations at the tyrosine or aspartic acid moieties in this sequence are incapable of catalyzing reverse transcription supports the critical role of this region.

An octapeptide conserved region located upstream of the noted tetrapeptide sequence contains invariant proline and glycine residues as well as the dipeptide Ser-Pro. The remaining four residues of the octapeptide sequence are characteristic of a group of viruses rather than common to all. By using this octapeptide sequence, the oncoviruses can be categorized into two groups and the lentiviruses and hepadnaviruses into two other groups (Fig. 1). Although only eight amino acids of the reverse transcriptase were compared, this classification is in general agreement with "phylogenetic trees" proposed by other investigators (for example, see ref. 24). A notable difference between classifying viruses solely by this region versus larger regions of the reverse transcriptase and endonuclease/integrase domains, as proposed by Sonigo *et al.* (17), is the suggestion that MuLV and FeLV may be more closely related to HTLV-I, HTLV-II, simian T-cell lymphotropic virus type I, and bovine leukemia virus (BLV) than previously appreciated.

Equally important, the 25 amino acids flanked by the "group-specific" and generic regions vary significantly among the retroviruses and hepadnaviruses and serve as a "fingerprint" for a specific retrovirus.

The limited size of the conserved region of the coding sequence for the reverse transcriptases precludes the use of direct hybridization analysis to identify related but as yet uncharacterized members of these viral families. Since the nucleic acid amplification procedure requires only short oligonucleotides as primers, we wished to determine if the PCR had a contribution to make in the identification of minimally similar viral genomes. The amplification procedure not only provides millionfold increases in the designated regions but also reduces the nucleic acid complexity by nearly the same factor. Therefore, specific hybridization with much shorter oligonucleotide probes than is typically feasible and decreased stringencies of hybridization are possible.

Due to the degeneracy of the translational code, invariant amino acids may vary in the coding nucleic acid sequences. We noted that the amino acids conserved within the group-specific and generic primer binding regions were encoded by a small number of possible codons. We investigated whether "degenerate" oligonucleotides (a mixture of oligonucleotides equal in size but containing various sequences that differ by at least one residue) would allow significant amplification relative to unique oligonucleotides for the PCR.

Several laboratories have identified endogenous viral genomes in the human genome that share significant nucleic acid sequence similarity with MuLV (21, 22) and mouse mammary tumor virus (26, 27). The region of reverse transcriptase discussed above is present in these endogenous genomes but generally contains greater diversity than the exogenous viral genomes in the generic primer binding region (see Fig. 1). Caution will have to be used if primers for this

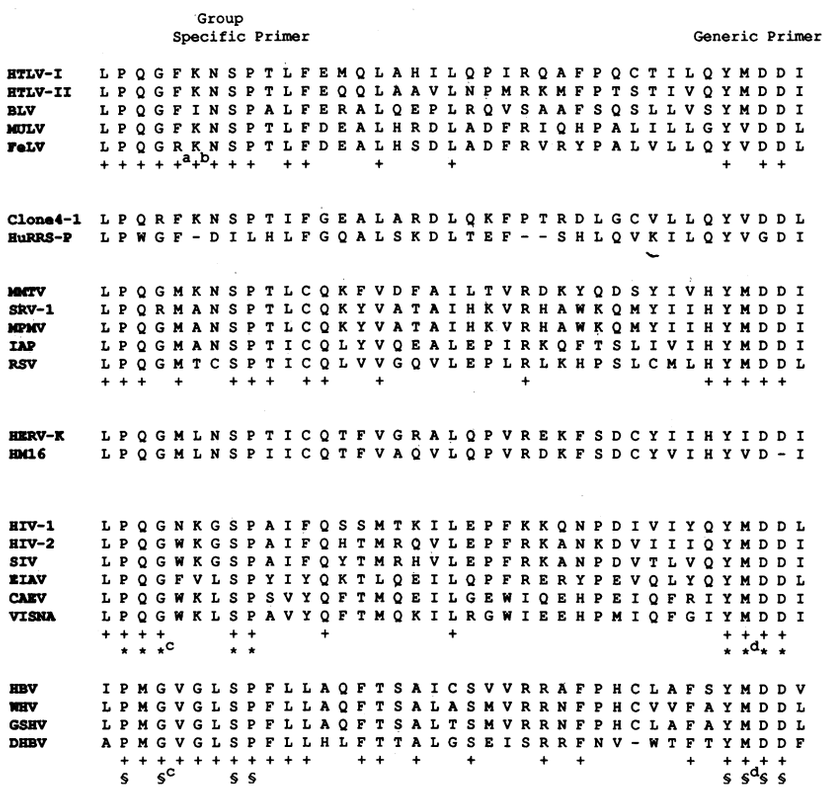

FIG. 1. Alignment of a conserved region of the reverse transcriptases of retroviruses and hepadnaviruses. The sequences for the aligned viruses have been reported and compared as follows: HTLV-I (17), HTLV-II (19), BLV (17), MuLV (17), FeLV (20), clone 4-1 (21), HuRRS-P (22), mouse mammary tumor virus [MMTV (19)], type D simian retrovirus [SRV-1 (23)], Mason-Pfizer monkey virus [MPMV (24)], Syrian hamster intracisternal A-particle gene [IAP (25)], Rous sarcoma virus [RSV (17)], Herv-k (26), HM16 (27), HIV-1 (17), HIV-2 (5), simian immunodeficiency virus [SIV (28)], equine infectious anemia virus [EIAV (19)], caprine arthritis encephalitis virus [CAEV (19)], visna (17), HBV (29), WHV (30), GSHV (31), and DHBV (32). The amino acid sequences abbreviated by the single-letter code are predicted from the nucleic acid sequences. Similarity of sequence in the octapeptide sequence labeled group-specific primer was used to establish the clustering of the viruses into groups. The highly conserved tetrapeptide sequence, Tyr-Met-Asp-Asp (Y M D D), is labeled as the generic region. +, Position of amino acids that are identical in all viruses in the group with the exception of FeLV ([a]), BLV ([b]), type D simian retrovirus (SRV-1) ([c]), MuLV ([d]), and FeLV ([d]). §, Positions of amino acids identical in all retroviruses and hepadnaviruses. Endogenous elements in the human genome similar to MuLV (clone 4-1 and HuRRS-P) and mouse mammary tumor virus (Herv-k and HM16) are included for comparison. Whenever possible, reports containing multiple sequences were chosen.

group of oncoviruses are to be used due to these closely related endogenous viral sequences.

Primer Selection. The detection of the members of the hepadnaviruses was chosen as a model system. Hepadnaviruses have been identified and characterized from humans (HBV) (33), woodchucks (WHV) (34), ground squirrels (ground squirrel hepatitis virus, GSHV) (35), and pekin ducks (DHBV) (36). Human isolates have been shown to vary by as much as 10% in nucleic acid sequence (37) whereas WHV and GSHV share ≈65% (31, 32) and DHBV shares 45% (32) nucleotide sequence similarity with HBV. Alignment of the codons of the hepadnaviral group-specific and generic primer binding sites as well as the intervening sequence is shown in Fig. 2. The strict conservation of amino acids and the presence of base alterations of the coding regions led us to test the use of degenerate oligonucleotides as primers for the PCR. Accordingly, the degenerate oligonucleotides synthesized included not only the base alterations of the known viruses but also all other codons for these amino acids.

The generic and group-specific primers contained 12 and 17 bases with 8 and 256 sequence isomers, respectively. This strategy approximates our approach to identify unclassified members of the hepadnaviruses. To simplify eventual cloning of the amplified products, 5'-terminal extensions of 9 nucleotides, which contained the recognition sequences for *Bam*HI and *Hin*dIII, were incorporated in the primers. Fig. 2 shows the sequence of the unique and degenerate primers synthesized for these studies. The degree of nucleic acid similarity of the region flanked by the primer binding regions (68%) and the decreased complexity of amplified DNA in the PCR suggested that a single oligonucleotide complementary to this region could be used to probe both the mammalian and avian members of the hepadnaviruses. Alternatively, long (38) and mixed-isomer (39, 40) oligonucleotide probes containing degenerate nucleic acid sequence predicted from the amino acid sequence of a specific protein could be used.

DNA Amplification. The unique primer pairs for WHV and DHBV as well as the degenerate primer pair for the hepadnaviruses were incorporated into DNA amplification reaction mixtures with genomic human DNA. The same DNA was used with approximately two copies per diploid genome of recombinant plasmid DNA containing either the WHV or DHBV viral genomes as target sequences. The amplified material was examined both by ethidium bromide/gel electrophoresis and Southern blot analysis to evaluate specific target amplification.

The ethidium bromide-stained patterns (Fig. 3*A*) suggested that the WHV and DHBV templates were amplified by the unique (lanes 1 and 5) and degenerate (lanes 3 and 6) primers. The generation of multiple nontarget bands of various molecular weights in PCRs (Fig. 3*A*, lanes 7–9) is a stochastic event that is primer-specific and occurs more frequently in the absence of template molecules complementary to the primers. The unique primer for WHV allows the detection of WHV sequences alone as denoted by the presence of the predicted 126-base-pair band (Fig. 3*B*). Human genomic DNA alone or with DHBV sequences added did not serve as a template for specific amplification (Fig. 3*B*, lanes 7 and 4, respectively). Correspondingly, the DHBV-specific primer pair allowed amplification of only the avian member of this class of viruses (Fig. 3*B*, lane 5). On the other hand, the degenerate oligonucleotide primers allowed the amplification of WHV and DHBV sequences (Fig. 3*B*, lanes 3 and 6). In separate experiments not shown here, DNA from recombinant plasmids and human liver biopsies containing the HBV genome have been shown to serve as templates for amplification with the unique mammalian and degenerate primer pairs. The similarity of sequence in this region for GSHV argues that this viral genome will serve as a template for amplification with these primers as well.

Sensitivity. Next, we determined the minimal number of infected cells among uninfected cells that could be detected by this procedure, assuming two copies of the viral genome per infected cell. The recombinant plasmid containing the WHV genome was added at 300,000, 30,000, 3000, 300, 30, and 3 copies per 150,000 cell-DNA equivalents (1 μg of human genomic DNA). Fig. 4, lane 5, clearly indicates that these primers will allow the detection of as few as 30 copies of viral genome equivalents per 150,000 cells.

Role of 5' Extensions on Primers. The dramatic specificity of the degenerate primers with linkers was surprising. Although only a fraction of the oligonucleotide primers were complementary to the template, efficient extension occurred at the elevated temperatures of the thermoresistant polymerase reaction. The 5' extensions of the oligonucleotide primers, which are not complementary to the template, are incorporated into the amplified products at the second cycle and all further cycles. As a result, we reasoned that more efficient extension by the primers on the newly synthesized

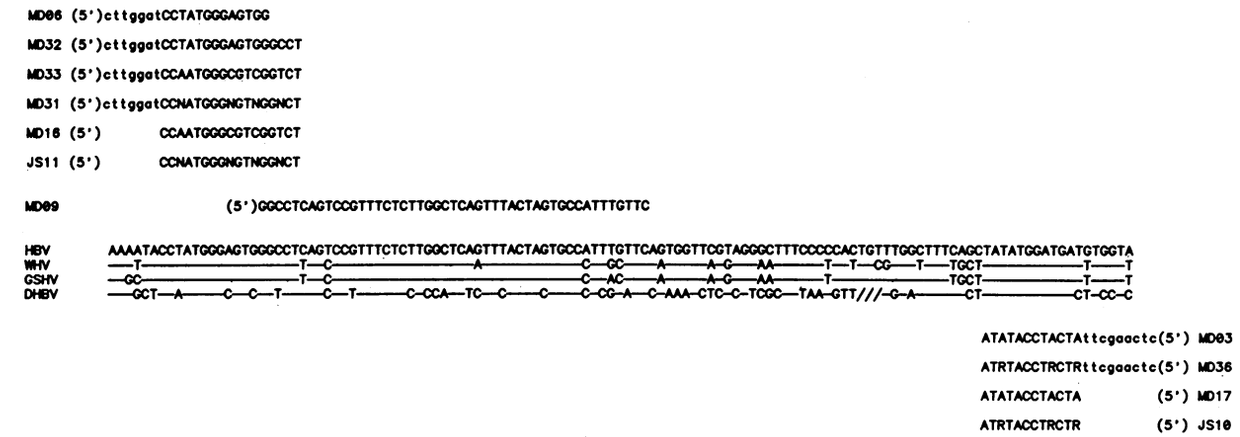

FIG. 2. Nucleotide sequence of the conserved region of the putative reverse transcriptase of the hepadnaviruses and oligonucleotide primers and probe. The sequence reported for HBV subtype adw2 (29) serves as the prototypical human hepadnavirus. Dashes represent identities in nucleotide sequence with the following aligned hepatitis viruses: WHV (30), GSHV (31), and DHBV (32). The oligonucleotide sequences of the primers and probe that appeared above and below the HBV sequence are homologous to the viral plus and minus strands, respectively. For clarity of the schematic PCR, the oligonucleotide sequences below the aligned hepadnaviral sequences are represented 3' to 5' rather than by the usual convention. The bases of the primer sequences denoted in lowercase letters represent 5' extensions that contain the recognition sequences for specific endonucleases (*Bam*HI and *Hin*dIII). The N and R indicate the four bases or the two purines, respectively, in that position.

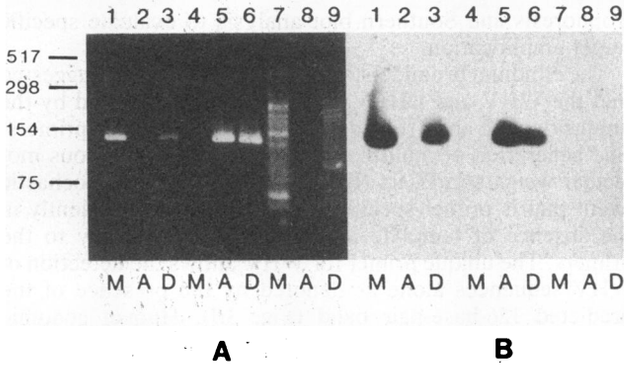

FIG. 3. Amplification of WHV and DHBV target sequences by using unique and degenerate oligonucleotide primers. The DNAs used for enzymatic amplification were 1 μg of an uninfected cell line (SC1) with 300,000 copies of a recombinant plasmid containing the entire genome of WHV (10) (lanes 1–3) or DHBV (11) (lanes 4–6) added or without plasmid (lanes 7–9). Thirty cycles of the PCR were carried out with mammalian-specific (MD32 and MD03; lanes M), avian-specific (MD33 and MD03; lanes A), or of degenerate (MD31 and 36; lanes D) oligonucleotides. The amplified products were examined by ethidium bromide/agarose gel electrophoresis (*A*) and Southern blot analysis (*B*) with MD09 as a probe.

amplified products would be accomplished due to the increased size of the priming duplex. We, therefore, compared the efficiency of amplification of the targeted viral sequences with primers lacking 5′ extensions. Although the predicted 126-base-pair product was readily observed in the ethidium bromide-stained gels by using primers with 5′ extensions, the amplified product was not detected if primers without additional sequences were incorporated into the amplification reaction (data not shown).

Detection of the amplified product is facilitated with a radioactively end-labeled probe oligonucleotide complementary to the sequence flanked by the primers, and the efficiency of the primers without 5′ extensions was lower by at least a factor of 25 by visual inspection of the autoradiogram (Fig. 5, compared lanes 1 and 2 with lanes 3 and 4). We, therefore, conclude that 5′ extensions of the primers play two important roles in the procedure. The additional sequences at the 5′ terminus of the primers increase the efficiency of the amplification process by increasing the stability of the priming oligonucleotide, and their incorporation permits the concomitant incorporation of endonuclease sites to simplify subsequent molecular cloning.

Cloning of Amplified DNA. Beyond detection, the next most important goal is to characterize the amplified product. By using the procedure described by Scharf *et al.* (13), the unique and degenerate primer pairs for DHBV resulted in plaque libraries with 92% and 88%, respectively, reverse

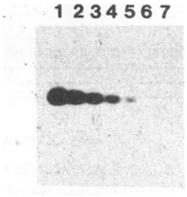

FIG. 4. Reconstitution studies with WHV target sequences for amplification. DNA from an uninfected human cell line (SC1) was amplified by 30 cycles of the PCR with the primers designated MD03 and MD06 in the absence or presence of added WHV genome. The DNAs used for enzymatic amplification were 1 μg of SC1 (the equivalent of 150,000 cells) alone (lane 7) and the same quantity of SC1 with 300,000 (lane 1), 30,000 (lane 2), 3000 (lane 3), 300 (lane 4), 30 (lane 5), and 3 (lane 6) copies of pWHV2 (10) added. Southern blot analysis was carried out as described (9).

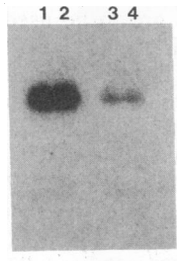

FIG. 5. Comparison of PCR primers with and without 5′ extensions. DNA from SC1 (1 μg) containing 300,000 copies of pWHV2 (10) was amplified for 30 cycles of the PCR with unique (MD32 and MD03) and degenerate (MD31 and MD36) primers containing 5′ extensions (lanes 1 and 2, respectively) and identical unique (MD16 and MD17) and degenerate (JS10 and JS11) oligonucleotides lacking 5′ extensions (lanes 3 and 4, respectively). The amplified products were examined by Southern blot analysis by using kinase-labeled MD09 as a probe.

transcriptase-specific candidate phages (data not shown). Thus, in addition to providing specific amplification of unique uncharacterized viral sequences, the rapidity of the molecular cloning and sequencing will greatly facilitate the study of viruses identified in this way.

DISCUSSION

The reverse transcriptases of retroviruses and hepadnaviruses share short, hyphenated regions of similarity (14). Alignment of the sequences of viral reverse transcriptases indicated that nucleic acid sequences with minimal degeneracy extending over only 12 and 16 bases encode the conserved amino acids of one particular region. The region examined in this report has several notable features. (*i*) A highly conserved, group-specific octapeptide is found at the extreme amino end. We have used this region of reverse transcriptase to classify the characterized exogenous and endogenous viral sequences into four categories. The oncoviruses are split into two distinct groups and the lentiviruses and hepadnaviruses represent the remaining two groups. It will be interesting to compare the analogous region of the spumaviruses to determine if they fall into yet a fifth category. (*ii*) The adjacent 25 amino acids found immediately downstream serve as a specific-virus fingerprint. (*iii*) A nearly invariant generic tetrapeptide sequence is found at the extreme carboxyl end.

We demonstrate here that this region can serve efficiently for DNA amplification directed by unique and degenerate oligonucleotides to the group-specific and generic primer binding sites. These data contribute to our understanding and utilization of the PCR. (*i*) At least one of the two primers for polymerase extension can be as short as 11 nucleotides. (*ii*) Degenerate oligonucleotides containing at least 256 specific isomers serve as primers for the exponential process. (*iii*) Noncomplementary 5′ extensions on short oligonucleotide primers dramatically increase the efficiency of amplification.

In contrast, a study with *Escherichia coli* DNA polymerase (13) demonstrated that 5′ modifications had no effect on the efficiency of amplification. Our interpretation of this difference is that, unlike our study, the conditions used by Scharf *et al.* (13) did not result in differential stability of the unmodified and 5′ modified primers. We also show that the resulting amplified products can be readily and efficiently cloned thereby simplifying sequence analysis. The use of degenerate primer pairs for exponential DNA amplification not only allows the detection of HBV, WHV, and DHBV but also may facilitate the detection of as yet uncharacterized hepadnaviruses in other organisms [for example, a suspected

virus in gray tree squirrels (41) and in Indian Palm tree squirrels†].

Although the studies reported here encompass exclusively the hepadnaviruses, we have designed similar primers and probes for the oncoviruses and lentiviruses that may prove fruitful in the ongoing studies in numerous laboratories to identify the possible viral agents involved in Kawasaki disease syndrome (42) and multiple sclerosis (43) or a whole host of other diseases speculated to have viral etiologies.

We also expect this procedure to detect and facilitate the identification of yet uncharacterized endogenous viral sequences. Depending on one's point of view, such endogenous viral sequences will represent either background noise in attempts to detect exogenous viruses or additional information about the structure of the eukaryotic genome.

Several investigators have identified short regions of amino acid sequence similarity in the polymerases of the RNA (44, 45) and DNA viruses (46). Perhaps, these regions or others yet to be identified may serve in a capacity similar to those described here for reverse transcriptases to assist in the search for as yet undescribed viruses involved in human disease. The use of this approach to develop generic diagnostic assays for batteries of pathogenic viruses has attractive potential.

Further, the explosion of available nucleic acid sequences encoding a myriad of proteins has led to the categorization of related proteins by short shared sequences. Perhaps, most notable are studies that have unmasked similarities among the protooncogenes. We suggest that coupling the DNA amplification procedure to the identification of highly conserved regions of known genes will provide a powerful tool to search for such genes.

†Raj Mehrota, K. G., International Symposium on Viral Hepatitis and Liver Disease, May 26–28, 1987, London, pp. 77A–78A (abstr.).

We thank C. Levenson, D. Spasic, and L. Goda for synthesis of oligonucleotides; S. Nilson and E. Ladner for preparation of the figures; and R. Bengelsdorf for excellent manuscript preparation. We also thank the Cetus PCR group for on-going criticism and support. We particularly would like to thank N. Arnheim, H. Erlich, D. Gelfand, M. Manos, F. McCormick, and T. White for critical reading of the manuscript. The research was supported by Cetus Corporation and Eastman Kodak.

1. Poiesz, B. J., Ruscetti, F. W., Gazdar, A. F., Bunn, P. A., Minna, J. D. & Gallo, R. C. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7415–7419.
2. Kalyanaraman, V. S., Sarngadharan, M. G., Robert-Guroff, M., Miyoshi, I., Blayney, D., Golde, D. & Gallo, R. C. (1982) *Science* **218**, 571–573.
3. Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W. & Montagnier, L. (1983) *Science* **220**, 868–871.
4. Gallo, R. C., Salahuddin, S. Z., Popovic, M., Shearer, G. M., Kaplan, M., Haynes, B. F., Palker, T. J., Redfield, R., Oleske, J., Safai, B., White, G., Foster, P. & Markham, P. D. (1984) *Science* **224**, 500–503.
5. Guyader, M., Emerman, M., Sonigo, P., Clavel, F., Montagnier, L. & Alizon, M. (1987) *Nature (London)* **326**, 662–669.
6. Blumberg, B. S., Gerstley, B. J., Hungerford, D. A., London, W. T. & Sutnick, A. I. (1967) *Ann. Intern. Med.* **66**, 924–931.
7. Prince, A. M. (1968) *Proc. Natl. Acad. Sci. USA* **60**, 814–821.
8. Mullis, K. B. & Faloona, F. A. (1987) *Methods Enzymol.* **155**, 335–350.
9. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuch, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–491.
10. Ogston, C. W., Jonak, G. J. Rogler, C. E., Astrin, S. M. & Summers, J. (1982) *Cell* **29**, 385–394.
11. Molnar-Kimber, K. L., Summers, J. & Mason, W. S. (1984) *J. Virol.* **51**, 181–191.
12. Vogelstein, B. & Gillespie, D. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 615–619.
13. Scharf, S. J., Horn, G. T. & Erlich, H. A. (1986) *Science* **233**, 1076–1078.
14. Toh, H., Hayashida, H. & Miyata, T. (1983) *Nature (London)* **305**, 827–829.
15. Kamer, G. & Argos, P. (1984) *Nucleic Acids Res.* **12**, 7269–7282.
16. Chiu, I., Callahan, R., Tronick, S. R., Schlom, J. & Aaronson, S. A. (1984) *Science* **223**, 364–370.
17. Sonigo, P., Alison, M., Staskus, K., Klatzmann, D., Cole, S., Danos, O., Retzel, E., Tiollais, P., Haase, A. & Wein-Hobson, S. (1985) *Cell* **42**, 369–382.
18. Larder, B. A., Purifoy, D. J. M., Powell, K. L. & Darby, G. (1987) *Nature (London)* **327**, 716–717.
19. Chiu, I.-M., Yaniv, A., Dahlberg, J. E., Gazit, A., Skuatz, S. F., Tronick, S. R. & Aaronson, A. (1985) *Nature (London)* **317**, 307–319.
20. Donahue, P. R., Hoover, E. A., Beltz, G. A., Riedel, N., Hirsch, V. M., Overbaugh, J. & Mullins, J. I. (1987) *J. Virol.* **62**, 722–731.
21. Repaske, R., Steele, P. E., O'Neill, R., Rabson, A. B. & Martin, M. A. (1985) *J. Virol.* **54**, 764–772.
22. Kroger, B. & Horak, I. (1987) *J. Virol.* **61**, 2071–2075.
23. Power, M. D., Marx, P. A., Bryant, M. L., Gardner, M. B., Barr, P. J. & Luciw, P. A. (1986) *Science* **231**, 1567–1572.
24. Sonigo, P., Barker, C., Hunter, E. & Wain-Hobson, S. (1986) *Cell* **45**, 375–385.
25. Ono, M., Toh, H., Miyata, T. & Awaya, T. (1985) *J. Virol.* **55**, 387–394.
26. Ono, M., Yasunaga, T., Miyata, T. & Ushikubo, H. (1986) *J. Virol.* **60**, 589–598.
27. Deen, K. C. & Sweet, R. W. (1986) *J. Virol.* **57**, 422–432.
28. Hirsch, V., Riedel, N. & Mullins, J. I. (1987) *Cell* **49**, 307–319.
29. Valenzuela, P., Quiroga, M., Zaldivar, J., Gray, P. & Rutter, W. J. (1980) in *Animal Virus Genetics*, eds. Fields, B., Jaenisch, R. & Fox, C. F. (Academic, New York), pp. 57–70.
30. Galibert, F., Chen, T. N. & Mandart, E. (1982) *J. Virol.* **41**, 51–65.
31. Seeger, C., Ganem, D. & Varmus, H. E. (1984) *J. Virol.* **51**, 367–375.
32. Mandart, E., Kay, A. & Galibert, F. (1984) *J. Virol.* **49**, 782–792.
33. Summers, J., O'Connel, A. & Millman, I. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4597–4601.
34. Summers, J., Smolec, J. & Snyder, R. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 4533–4537.
35. Marion, P. L., Oshiro, L. S., Regnery, D. C., Scullard, G. H. & Robinson, W. S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2941–2945.
36. Mason, W. S., Seal, G. & Summers, J. (1980) *J. Virol.* **36**, 829–836.
37. Tiollais, P., Charnay, P. & Vyas, G. N. (1981) *Science* **213**, 406–411.
38. Anderson, S. & Kingston, I. B. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6863–6842.
39. Suggs, S. V., Wallace, R. B., Hirose, T., Kawashima, E. H. & Itakura, K. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6613–6617.
40. Chardin, P. & Tavitian, A. (1986) *EMBO J.* **5**, 2203–2208.
41. Feitelson, M. A., Millman, I., Halbherr, T., Simmons, H. & Blumberg, B. S. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2233–2237.
42. Burns, J. C., Geha, R. S., Schneeberger, E. E., Newburger, J. W., Rosen, F. S., Glezen, L. S., Huang, A. S., Natale, J. & Leung, D. Y. M. (1986) *Nature (London)* **323**, 814–816.
43. Koprowski, H., DeFreitas, E. C., Harper, M. E., Sandeberg-Wollheim, M., Sheremata, W. A., Robert-Guroff, M., Sazinger, C. W., Feinberg, M. B., Wong-Staal, F. & Gallo, R. C. (1985) *Nature (London)* **318**, 154–160.
44. Franssen, H., Leunissen, J., Goldbach, R., Lomonossoff, G. & Zimmern, D. (1984) *EMBO J.* **3**, 855–861.
45. Argos, P., Tucker, A. D. & Philipson, L. (1986) *Virology* **149**, 208–216.
46. Larder, B. A., Kemp, S. D. & Darby, G. (1987) *EMBO J.* **6**, 169–175.