Short Technical Reports

ondary structure specific to the mutant sequence was modeled. However, it was not drastically different from the wild-type sequence and likely does not affect the efficiency of PCR amplification (data not shown).

In an attempt to disrupt any secondary structure specific to the mutant sample, we designed four downstream primers identical to D3 to contain unique mismatches. In theory, the mismatches could destabilize a hairpin or other structure in the template, enabling the mutant sample to be amplified. These mismatched primers were paired with U1 primer for PCR amplification. Inhibition was observed for the mutant samples, while efficient amplification occurred for the wild-type samples (data not shown). These experiments suggest that the inhibition may be due to factors unrelated to template secondary structure.

The experiments above have demonstrated the inhibition of PCR amplification by a point mutation in the factor IX gene. Although the prevalence of such a dramatic inhibition may not be high, the incidence of more subtle differences in the efficiency of amplification of two samples with seemingly insignificant differences in DNA sequence may be significantly greater. This phenomenon should be considered when performing PCR-based methods that critically depend on the efficiency of PCR amplification.

REFERENCES

1. Stoflet, E.S., D.D. Koeberl, G. Sarkar and S.S. Sommer. 1988. Genomic amplification with transcript sequencing. *Science* 239:491-494.
2. Walsh, P.S., H.A. Erlich and R. Higuchi. 1992. Preferential PCR amplification of alleles: mechanisms and solutions. *PCR Methods Appl.* 1:241-250.
3. Yoshitake, S., B.G. Schach, D.C. Foster, E.W. Davie and K. Kurachi. 1985. Nucleotide sequence of the gene for human factor IX (anti-hemophilic factor B). *Biochemistry* 24:3736-3750.

We thank Marquis Gacy and Cynthia McMurray, Ph.D., for providing the values for FOLDDNA energies that were used with FOLDRNA, Mr. James Prescott for his help in quantitative analysis with the PhosphorImager and Mary Johnson for expert secre-

tarial assistance. The work was supported by RO1 MH44276. Address correspondence to Steve S. Sommer, Department of Biochemistry and Molecular Biology, Mayo Clinic/Foundation, Rochester, MN 55905, USA.

Received 28 March 1996; accepted 29 July 1996.

Qiang Liu, Erik C. Thorland and Steve S. Sommer
*Mayo Clinic/Foundation
Rochester, MN, USA*

Simple Method for Selective Amplification of cDNA from a Defined Promoter

BioTechniques 22:300-306 (February 1997)

ABSTRACT

A simplified technique for the detection of transcripts from a defined promoter is described. After reverse transcription, a PCR target sequence is selectively added to the 3' end of cDNA strands by DNA polymerase extension directed by an oligonucleotide template. Those cDNA molecules that do not have ends within a few nucleotides of the promoter start site are not extended and thus are excluded from subsequent amplification. Even when amplified products are visualized by ethidium bromide staining of agarose gels, this method requires only 1% of the RNA usually needed for detection of mRNA by standard RNase protection utilizing radiolabeled probes. In contrast to direct detection of cDNA by PCR, this procedure restricts amplification to a narrow subset of transcripts even when other overlapping colinear transcripts are present. We call this detection procedure specific amplification of cDNA ends (SPACE).

INTRODUCTION

While studying gene regulation in papilloma viruses, we sought a sensitive method for detecting transcripts initiating from a specific promoter whose start site may overlap with transcripts from adjacent promoters. To detect such transcripts, primer extension or RNase protection assays have been used with limited sensitivity. We describe here a polymerase chain reaction (PCR)-based amplification strategy that detects only cDNA molecules with defined ends. In contrast to the cDNA end amplification methods previously described [e.g., rapid amplification of cDNA ends (RACE) reviewed in Reference 3], this method exhibits specificity and is therefore referred to as specific amplification of cDNA ends (SPACE).

MATERIALS AND METHODS

Figure 1 diagrammatically depicts

Short Technical Reports

oligonucleotides used in the SPACE protocol. The extension oligonucleotide has the sequence ggcgtacaaccctcGAACTGCAATGTTTCAGGAC-CCgttt; where the uppercase letters correspond to human papilloma virus (HPV)-16 sequence nucleotides (nt) 96–117, and the lowercase letters represent non-base-pairing nucleotides. The standard reaction was performed in 50 μ L and contained cDNA or DNA targets, 5 pmol extension primer, 200 μ M each of dATP, dTTP, dGTP and dCTP, 1 $\times$ PCR buffer and 1 U of *Taq* DNA Polymerase (Boehringer Mannheim, Indianapolis, IN, USA). After 4 cycles of 30 s each at 94° and 50°C, and 1 min at 72°C, 50 pmol each of the amplification primer (ggcgtacaaccctcGAAC) and downstream primer (GTCCAGATGTCTTTGCTTTTCTTCAGGACACAGTGGCT; HPV-16 nt 463–426) were added, and 35 cycles of 30 s each at 94° and 55°C, and 1 min at

72°C were carried out.

Transfection of HeLa cells with pSV2N-*cat*, RNA preparation and RNase protection analysis was performed as previously described (5). The HPV-16 P97 RNase protection probe was synthesized with [α -³²P]-rUTP by T7 transcription of a PCR-generated DNA template to give the antisense strand extending from HPV-16 nt 307 to nt 45. For synthesis of cDNA, 1 μ g total RNA dissolved in diethyl pyrocarbonate (DEPC)-treated H₂O in a volume of 10.5 μ L was mixed with 1 μ L random hexamer oligonucleotides (20 μ M; CLONTECH Laboratories, Palo Alto, CA, USA), heated at 70°C for 5 min and cooled on ice. The final reaction mixture of 20 μ L, which also contained 500 μ M of each deoxyribonucleoside triphosphate, 10 mM dithiothreitol, 200 U Moloney murine leukemia virus (M-MLV) reverse transcriptase and buffer (Life Technologies, Gaithersburg, MD,

USA) were incubated at 42°C for 1 h. After 5 min at 94°C to inactivate enzymes, the cDNA was diluted to 100

Figure 1. Scheme for specific detection of transcripts. The arrows indicate the 3'-ends of single oligonucleotide strands. See Results and Discussion for discussion.

Short Technical Reports

μ L with DEPC-treated H_2O . The DNA templates containing defined ends (see Figure 2A) were synthesized by PCR using primers with the indicated ends and *Pfu* DNA Polymerase (Stratagene, La Jolla, CA, USA), which produces a high yield of blunt-ended amplified DNA (2). PCR to amplify a 263-nt fragment from *cat* cDNA used the primers GCGCAAGCTTTTCAGGAGCTAAGGAAGCTAA and CCTAATACGACTCACTATAGGGAGAGCATTCATCAGGCGGGCA with 35 cycles of 30 s each at 94° and 55°C, and 1 min at 72°C.

RESULTS AND DISCUSSION

To amplify a specific mRNA when other overlapping mRNAs may be present, the method outlined in Figure 1 was used. For a model system, the major early promoter of HPV-16, P97, which was designated transcript A, was selected. In the diagram, an arrowhead indicates the 3' end. The 3' end of the cDNA generated from transcript A is the anti-sequence of the transcript start site. In **step 1**, the extension oligonucleotide is annealed to the 3' terminal 22

nt of the cDNA. The oligonucleotide's 15 5'-bases are dissimilar to the HPV-16 DNA upstream of nt 97 (designated by the box in Figure 1), and its 3' end also contains a 4-nt, non-annealing sequence (designated by the smaller box). In **step 2**, the cDNA is 3'-extended by *Taq* DNA polymerase and incorporates the dissimilar anti-sequence of the extension oligonucleotide. The oligonucleotide, however, does not extend from its 3' end because it does not base-pair-match the cDNA. In **step 3**, the amplification primer, which anneals perfectly to the newly synthesized 3' end of the cDNA created in **step 2**, is extended to form a complement to the cDNA. This new double-stranded DNA acts as a template in **step 4** for conventional PCR amplification directed by the amplification primer and the downstream primer.

The cDNA produced from a colinear mRNA (transcript B) that initiated upstream of the start site of transcript A would also anneal to the extension oligonucleotide. This complex would not serve as a polymerase template because neither 3' end would contain correct base-pair matching. The anti-sequence for the amplification primer would not be incorporated into this DNA, and PCR amplification could not proceed. Thus, only transcripts initiating very near to nt 97 should be ampli-

fied by this scheme. Furthermore, contaminating DNA will not be amplified, since it will not precisely contain an end at nt 97.

To test for the ability of the SPACE protocol to discriminate between cDNAs with ends near to the P97 start site, amplifications were performed with the control DNA templates outlined in Figure 2A. No amplification was seen with 10^6 copies of either the plasmid (template *a*), which contains no ends, or the fragment (template *b*) extending to nt 94 (Figure 2B, lanes 1 and 2). Efficient amplification with the SPACE protocol of the predicted 236-nt band resulted only for the DNA (template *c*) with a 5' end at nt 97 (Figure 2B, lane 3). With as few as 1000 copies of this input template DNA, a faint band was detectable (data not shown). Although no amplification was observed under the conditions shown for the DNA target (template *d*) with a 5' end at nt 104 (Figure 2B, lane 4), if the annealing temperature were lower than 50°C for the initial extension step or if more than 10^6 copies of DNA were added, a signal was detected (data not shown). In standard amplification reactions with template *c*, where either the extension oligonucleotide or the amplification primer were omitted, no product was observed (Figure 2B, lanes

Figure 2. Specific amplification of DNAs with defined ends. (A) DNA targets used in amplification reactions. Templates *b–d* are derived from pHPV-16, a plasmid containing the entire HPV-16 coding sequence (target *a*), by PCR using *Pfu* DNA polymerase and primers that encompass the indicated nucleotide coordinates. (B) Amplified DNAs were detected by ethidium bromide staining after agarose gel electrophoresis. Amplification used the SPACE protocol with one million molecules of indicated targets (lanes 1–6) in each reaction. Amplification primer or extension primer was omitted in lane 5 or 6, respectively.

Figure 3. Amplification of specific mRNAs. (A) Amplified DNAs were detected by ethidium bromide staining after agarose gel electrophoresis. Amplification targets were cDNA-generated from RNA extracted from HeLa cultures 24 h after transfection with pSV2N-*cat* (lanes 1 and 2) or from SiHa cells (lanes 3–6). Amplification with the SPACE protocol directed at HPV-16 P97 was used in lanes 2–7. Conventional PCR using primers specific for the *cat* gene was used in lane 1. The cDNA derived from 100 ng RNA was used in lanes 1, 2 and 7, while 5, 10, 25 and 50 ng were used in lanes 3–6, respectively. The predicted position of migration for the spliced and unspliced P97 as well as *cat*-derived amplified DNAs are indicated. (B) Fragments protected from RNase digestion by hybridization with 20 μ g SiHa RNA (lane 1) using an antisense ^{32}P -labeled riboprobe spanning the P97 promoter of HPV-16 and separated by denaturing acrylamide gel electrophoresis (lane 2) were detected by autoradiography. Migration position of undigested probe and fragments protected by spliced and unspliced P97 transcripts are indicated.

Short Technical Reports

5 and 6). The protocol therefore selectively amplifies templates with ends in a narrow range.

To test for efficient amplification of P97 cDNA by the SPACE protocol, we used reverse-transcribed total SiHa RNA as a target. SiHa is a cell line derived from a cervical carcinoma that contains a few copies of integrated HPV-16 DNA. The P97 promoter of HPV-16 (4) is active in these cells (1). For a negative control, we used total RNA extracted from HeLa cells that had been transfected with pSV2N-*cat* and contains *cat* mRNA detectable by RNase protection (data not shown). After reverse transcription, the cDNA can be amplified using *cat*-specific primers (Figure 3A, lane 1). The HeLa cell line contains transcriptionally active DNA derived from HPV-18, a divergent papilloma virus type. This cDNA should not contain sequences closely related to HPV-16. As predicted, no amplification with the SPACE protocol was seen (Figure 3A, lane 2). However, cDNA derived from as little as 10 ng of SiHa RNA gave detectable amplification products (Figure 3A, lanes 3–7). As expected, two bands are seen, corresponding to the unspliced and spliced P97 derived transcripts (4). In comparison, 20 µg of this RNA used to protect an antisense riboprobe (Figure 3B, lane 1) spanning the P97 promoter from RNase digestion gave detectable bands corresponding to spliced and unspliced P97 mRNA after 8 h of autoradiography (Figure 3B, lane 2). If autoradiography is extended to one week, the practical limit of detection in this RNase protection assay is about 1 µg RNA. Thus, approximately 1/100 of the RNA is required for detection of P97 transcripts by the SPACE protocol as by the RNase protection assay.

We believe the SPACE method has some unique properties that make it a useful technique for detection of transcripts. Since the SPACE protocol is able to discriminate between cDNAs containing different ends, it allows detection of transcripts with a specific start site. Compared to other assays that detect specific RNA ends, such as S1 or RNase protection assays, amplification techniques are more sensitive. Even with detection of amplified products by ethidium bromide staining of agarose

gels, the SPACE method requires 1/100 of the input RNA compared with RNase protection for detection of transcripts. With ³²P-based detection, an additional one to two orders of magnitude increase in sensitivity would be expected.

The RACE technique depends on extending the 3' end of the cDNAs by nonspecific enzymatic extension or ligation of an oligonucleotide to provide a target sequence for an amplification primer. Thus, the RACE technique (reviewed in Reference 3) can define unknown mRNA 5' ends and is often used to clone full-length cDNAs when a partial sequence is known. RACE has not been widely used to measure specific transcripts originating at different transcription start sites in a gene since any transcript initiating within several hundred or more nucleotides would be amplified. When the transcript of interest initiates within the body of an abundant colinear transcript, a RACE protocol would preferentially amplify the more prevalent message. In contrast, mRNAs amplified by the SPACE protocol are restricted to those with predefined 5' ends. Therefore, detection of low-level mRNAs is not inhibited by more prevalent colinear transcripts. An additional advantage of the SPACE protocol is that no ligation steps are required. Because fewer steps are involved, this procedure should be more reproducible for use in semiquantitative competitive assays using a synthetic internal standard RNA with the appropriate 5' end.

REFERENCES

1. Baker, C.C., W.C. Phelps, V. Lindgren, M.J. Braun, M.A. Gonda and P.M. Howley. 1987. Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. *J. Virol.* 61:962-971.
2. Hu, G. 1993. DNA polymerase-catalyzed addition of nontemplated extra nucleotides to the 3' end of a DNA fragment. *DNA Cell Biol.* 12:763-770.
3. Schaefer, B.C. 1995. Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. [Review]. *Anal. Biochem.* 227:255-273.
4. Smotkin, D., H. Prokoph and F.O. Wettstein. 1989. Oncogenic and nononcogenic human genital papillomaviruses generate the E7 mRNA by different mechanisms. *J. Virol.* 63:1441-1447.
5. Ushikai, M., M.J. Lace, Y. Yamakawa, M. Kono, J. Anson, T. Ishiji, S. Parkkinen, N.

Wicker, M.-E. Valentine, I. Davidson, L.P. Turek and T.H. Haugen. 1994. *trans* activation by the full-length E2 proteins of human papillomavirus type 16 and bovine papillomavirus type 1 in vitro and in vivo: cooperation with activation domains of cellular transcription factors. *J. Virol.* 68:6655-6666.

L.P.T. and T.H.H. receive funding from the Department of Veterans Affairs. L.P.T. also receives funding from the American Cancer Society (VM-63). We thank Jim Anson for performing the RNase protection assay. Address correspondence to Thomas H. Haugen, Department of Pathology, University of Iowa College of Medicine, Iowa City, IA 52242, USA. Internet: thomas-haugen@uiowa.edu

Received 16 April 1996; accepted 2 August 1996.

Pavel Bostik, Lubomir P. Turek and Thomas H. Haugen
*Department of Veterans Affairs
Medical Center and
University of Iowa
Iowa City, IA, USA*