antibody fragments. Curr. Opin. Struct. Biol. *12*:503-508.
17. **Smith, G.P.** 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science *228*: 1315-1317.
18. **Williamson, R.A., R. Burioni, P.P. Sanna, L.J. Partridge, C.F. Barbas III, and D.R. Burton.** 1993. Human monoclonal antibodies against a plethora of viral pathogens from single combinatorial libraries. Proc. Natl. Acad. Sci. USA *90*:4141-4145.
19. **Barbas, C.F., III, J.D. Bain, D.M. Hoekstra, and R.A. Lerner.** 1992. Semisynthetic combinatorial antibody libraries: a chemical solution to the diversity problem. Proc. Natl. Acad. Sci. USA *89*:4457-4461.
20. **QuikChange® Multi Site-Directed Mutagenesis Kit.** Stratagene Co., 11011 N. Torrey Pines Rd., La Jolla, CA 92037, U.S. patent Nos. 5,789,166; 5,923,419; 6,391,548; 6,183,997; 5,948,663; 5,866,395; 5,545,552; and patents pending.
21. **Hogrefe, H.H., J. Cline, G.L. Youngblood, and R.M. Allen.** 2002. Creating randomized amino acid libraries with the QuikChange® Multi Site-Directed Mutagenesis Kit. BioTechniques *33*:1158-1165.
22. **Momany, C., L.C. Kovari, A.J. Prongay, W. Keller, R.K. Gitti, B.M. Lee, A.E. Gorbalenya, L. Tong, et al.** 1996. Crystal structure of dimeric HIV-1 capsid protein. Nat. Struct. Biol. *3*:763-770.
23. **Barbas, C.F., III, D.R. Burton, J. K. Scott, and G.J. Silverman (Eds.).** 2001. Phage Display: A Laboratory Manual. CSH Laboratory Press, Cold Spring Harbor, NY.
24. **Engvall, E., K. Jonsson, and P. Perlmann.** 1971. Enzyme-linked immunosorbent assay. II. Quantitative assay of protein antigen, immunoglobulin G, by means of enzyme-labelled antigen and antibody-coated tubes. Biochim. Biophys. Acta *251*:427-434.

Received 21 May 2003; accepted 27 June 2003.

Address correspondence to Cory Momany, Room 372 Wilson Pharmacy Building, College of Pharmacy, University of Georgia, Athens, GA 30602, USA. e-mail: cmomany@mail.rx.uga.edu

Random priming PCR strategy to amplify and clone trace amounts of DNA

Nianxiang Zou, Susan Ditty, Bingjie Li, and Shyh-Ching Lo
Armed Forces Institute of Pathology, Washington, DC, USA

BioTechniques 35:758-765 (October 2003)

Here we report a new methodology to study trace amounts of DNA of unknown sequence using a two-step PCR strategy to amplify and clone target DNA. The first PCR is carried out with a partial random primer comprised of a specific 21-nucleotide 5′ sequence, a random heptamer, and a 3′ TGGC clamp. The second PCR is carried out with a single 19-nucleotide primer that matches the specific 5′ sequence of the partial random primer. Using human and Mycoplasma genitalium *DNA as examples, we demonstrated the efficiency of this approach by effectively cloning target DNA fragments from 1 pg DNA sample. The cloning sensitivity could reach 100 fg target DNA templates. Compared to the strategy of first adding adapter sequences to facilitate the PCR amplification of unknown sequences, this approach has the advantage of allowing for the amplification of DNA samples in both natural and denatured forms, which provides greater flexibility in sample preparation. This is an efficient strategy to retrieve sequences from trace DNA samples from various sources.*

INTRODUCTION

DNA cloning is one of the most commonly used methods in molecular biology. The technique allows one to obtain a large quantity of specific DNA molecules for study or characterization. However, there is a limitation to this powerful technique. In conventional DNA cloning, a significant amount of target molecules is needed to ensure efficient ligation with vectors. With PCR, cloning trace amounts of DNA can be achieved by first amplifying the target DNA, given that partial sequences are known for the DNA segment of interest and that PCR primers can be designed. However, cloning trace amounts of DNA with unknown sequence presents a difficult task in many situations. Conventional or regular PCR-based cloning strategies are typically unable to fulfill this task.

Many amplification strategies using small amounts of DNA or RNA with unknown sequence have been reported (1–5). Random primers tagged by a specific sequence at the 5′ end were commonly used to introduce primer sequences to facilitate PCR amplification (4,5). Most of these applications still require samples in nanogram amounts (4,5). To attain higher sensitivity, sequence-independent single-primer amplification (SISPA) was developed to amplify unknown target DNA molecules with PCR and has been shown to facilitate cloning from 1 pg DNA after amplification (3). SISPA first requires ligation of an adapter onto the target population of blunt-ended DNA molecules. The resulting DNA with common end sequence allows for PCR amplification by a single primer made from the adapter sequence.

Random priming strategies generating single-stranded DNA (ssDNA) molecules are frequently used to randomly multiply DNA or synthesize cDNA from RNA molecules (6). The most common applications include probe labeling, multiplying DNA or RNA from small samples, and generating random in vitro mutagenesis and recombination (1,2,7–10). To clone trace amounts of unknown DNA from various sources of different physical and biological properties, we have developed a two-step PCR method that is partially based on random priming strategy. The first PCR uses a single primer composed of seven random nucleotides flanked by a specific 5′ tag and a 3′ clamp (Figure 1, Ran5-29). A *Pst*I site (New England Biolabs, Beverly, MA, USA) is included in the 5′ tag. The internal seven random nucleotides and the 3′ clamp TGGC enable the primer to anneal with any seven nucleotides with GCCA immediately upstream. The TGGC clamp increases the primer's 3′-specific annealing from 7 to 11 nucleotides. The second PCR is performed using a single specific

primer (Fix5-29) matching 19 nucleotides of the 5′ tag sequence of Ran5-29. The second PCR markedly amplifies the DNA strands synthesized in the first PCR (Figure 1). The *Pst*I site introduced at both ends of the PCR products facilitates cloning of the PCR products into the vector, which is linearized by the same restriction enzyme.

To demonstrate the feasibility of this cloning strategy, we used DNA preparations from human and *Mycoplasma genitalium*, which represents one of the most complex and one of the simplest genomes, respectively. Our results showed that with both DNA samples, this strategy could reach a cloning sensitivity of 100 fg of target DNA samples.

MATERIALS AND METHODS

DNA Preparation

Human peripheral blood DNA and *M. genitalium* DNA were prepared using the DNeasy® Tissue Extraction Kit (Qiagen, Valencia, CA, USA). The DNA samples were quantitated using a MBA 2000 Spectrometer (PerkinElmer Life Sciences, Gaithersburg, MD, USA).

Random and Specific Primer PCRs

The partially random primer Ran5-29 (5′-GTTCTACACGAGTCACTG-CAGNNNNNNNTGGC-3′) was used in the first PCR. The reactions were performed in 50-µL volumes containing sample DNA, 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 200 µM each dNTP, 0.5 µM primer, and 0.5 U AmpliTaq® Gold DNA Polymerase (PerkinElmer) in a GeneAmp® 9600 Automated Thermal Cycler (PerkinElmer). PCR conditions were 95°C for 10 min, followed by 30 cycles of 95°C for 1 min, 58°C for 1 min, 72°C for 1 min, and a final 5-min extension step at 72°C. The primer Fix5-29 (5′-GTTCTACACGAGTCACTGC-3′) was used in the second PCR. The reaction was performed under the same conditions as the first PCR, except that 55°C was used for annealing and there were 45 cycles. The products of the first PCR were purified with the GENECLEAN® Kit (Bio101, Carlsbad, CA, USA). One quarter of each purified product was used as the template for the second PCR with the specific primer. To minimize the background DNA level, which may interfere with cloning efficiency at very low target DNA amounts, the preparation and dilution of target DNA samples were handled carefully, and reaction mixtures were prepared in a PCR workstation (Misonix, Farmingdale, NY, USA).

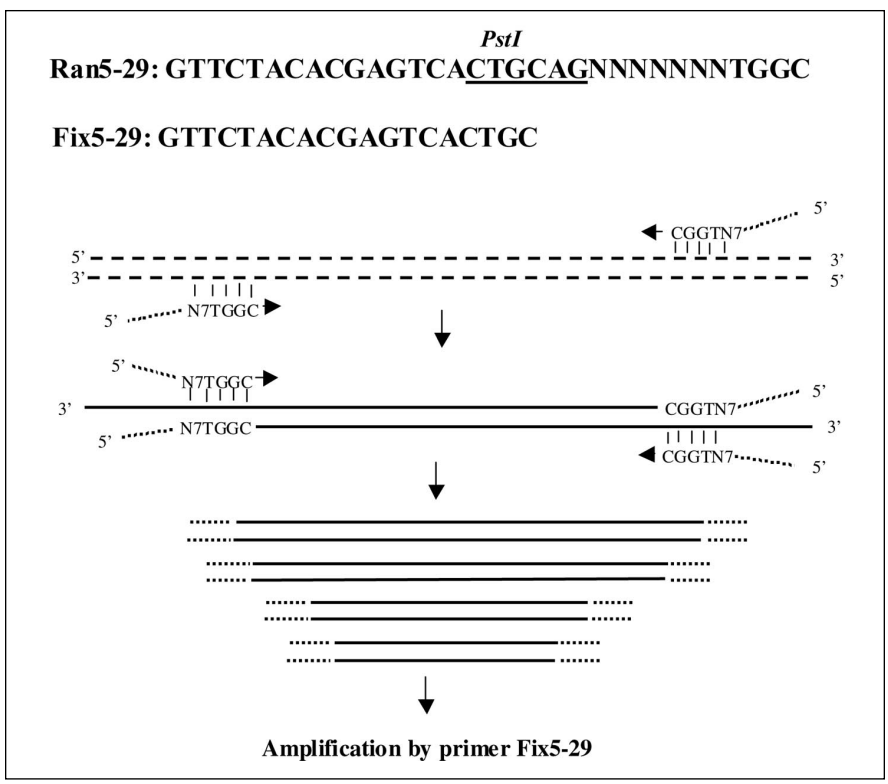

Figure 1. Random and specific PCR strategy. Upper panel, partially random primer and specific primer sequences. Lower panel, PCR amplification process. The broken lines indicate templates from trace amount of input DNA. The solid lines indicate newly generated DNA strands. The arrows indicate the direction of DNA synthesis, and the dotted lines indicate the specific sequence of the partially random primer.

DNA Cloning

Target DNA samples and the vector pUC19 were digested with *Pst*I. The linearized vector was treated with calf intestine alkaline phosphatase (CIAP; New England Biolabs) to remove the 3′ phosphate group to prevent vector self-ligation. Ligation reactions were performed in 5-µL volumes containing target DNA, 0.1 µg vector, 0.5 U T4 DNA Ligase, and 1× T4 DNA Ligase Buffer (both from New England Biolabs), and incubated at 14°C overnight. The ligation mixture was used to transform Library Efficiency® *Escherichia coli* DH5α Competent Cells (Invitrogen, Carlsbad, CA, USA), and the transformed *E. coli* cells were plated on LB agar plates containing 50 µg/mL ampicillin, 40 µg/mL isopropylthio-β-D-galactoside (IPTG), and 40 µg/mL X-gal. The plates were cultured at 37°C overnight.

Colony Screening and DNA Sequencing

E. coli colonies harboring recombinant plasmids were screened by blue and white color selection, followed by PCR with M13 forward and reverse primers. PCR was performed in 20-µL volumes containing 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 2.5 mM $MgCl_2$, 200 µM each dNTP, 0.5 µM each primer, and 0.2 U AmpliTaq DNA polymerase. A white *E. coli* colony was inoculated into the PCR mixture as template. PCR conditions were 30 cycles of 95°C for 1 min, 55°C for 1 min, 72°C for 1 min, and a final 5-min extension step at 72°C. PCR products were gel-purified. Nucleotide sequences were determined using the flanking M13 primers using the ABI PRISM® BigDye™ Terminator Cycle Sequencing Version 2.0 Ready Reaction Kit on

Table 1. Sequence Information of the Cloned DNA Fragments

Human DNA (pg)	Clone No.	N7TGGC[a]	Insert Size (bp)	BLASTN Results	Myc DNA	Clone No.	N7TGGC[a]	Insert Size (bp)	BLASTN Results
100 pg	1	-	>791	Human	100 pg	1	+	>703	Mycoplasma
	2	- -	310	Human		2	+	>635	Mycoplasma
	3	- -	691	Human		3	- +	720	Mycoplasma
	4	- -	291	Human		4	+	>470	Mycoplasma
	5	- -	234	Human		5	+ +	746	Mycoplasma
10 pg	1	-	359	Human	10 pg	1	+	>620	Mycoplasma
	2	- +	167	Human		2	+	>627	Mycoplasma
	3	- +	566	Human		3	-	>733	Mycoplasma
	4	- -	445	Human		4	+	648	Mycoplasma
	5	- +	566	Human		5	+ +	710	Mycoplasma
1 pg	1	- -	306	Human	1 pg	1	+ +	656	Mycoplasma
	2	- -	306	Human, same as 1		2	+	>758	Mycoplasma
	3	- -	258	Human		3	- -	706	no good match
	4	- -	306	Human, same as 1		4	- -	603	Mycoplasma
	5	- -	258	Human, same as 3		5	+ -	693	Mycoplasma
100 fg	1	- -	158	no good match	100 fg	1	- -	375	no good match
	2	- -	261	Human		2	- -	550	Mycoplasma
	3	- -	261	Human, same as 2		3	- -	375	no good match
	4	- -	401	Human		4	+ +	490	no good match
10 fg	1	- -	249	*Caulobacter crescentus* (85%)	10 fg	1	- -	375	no good match
	2	- -	143	*Mesorhizobium loti* (84%)		2	- -	384	*Burkholderia cepacia* (82%)
						3	- -	259	no good match

[a]The symbols "+" and "-" indicate the presence or absence, respectively, of the N7TGGC sequence marker at the 3′ or 5′ ends of each insert confirmed by sequencing. Each cloned DNA fragment was sequenced from one terminus. The presence or absence of the N7TGGC marker could only be confirmed on one end of some long cloned DNA fragments.

an ABI PRISM 377 Automated DNA Sequencer (Applied Biosystems, Foster City, CA, USA).

RESULTS AND DISCUSSION

To test the efficiency of our cloning strategy using our partially random (Ran5-29) and specific (Fix5-29) primers, human and *M. genitalium* DNA quantities of 1 ng, 100, 10, and 1 pg, and 100 and 10 fg were used in the Ran5-29 PCRs. Faint DNA smears were produced for each sample including double-distilled water (data not shown). Higher amounts of DNA produced slightly stronger smear signals, but dramatic differences were not observed. The smear produced by the double-distilled water as template was apparently amplified from the background DNA in our PCR systems.

The annealing temperature appeared to be important for the Ran5-29 reaction. We tested 45°, 55°, and 58°C by observing smear signals on agarose gels that were produced by the subsequent Fix5-29 PCR. At 45°C, no smear signals of the PCR products were obtained. PCR with an annealing temperature of 58°C generated more consistent smears than PCR with an annealing temperature of 55°C. As with high primer concentrations, at low temperatures, the random primer could anneal with many areas of the template with lower stringency, which favors the generation of short DNA chains (11). At high temperatures, which is a more stringent annealing condition, primer annealing is less efficient. However,

once DNA synthesis initiates, it tends to generate longer chains because the primer annealing sites are more widely spread on the template. To be amplified by primer Fix5-29, the primer sequence needs to be present at both termini of the dsDNA molecules (Figure 1).

We found that 30 cycles was optimum for Ran5-29 PCR. Ran5-29 PCR at 15 or 25 cycles significantly reduced the production of smear signals in subsequent Fix5-29 PCR and thus reduced the efficiency of obtaining recombinant plasmids in the steps that followed. On the other hand, 45 cycles of Ran5-29 PCR produced slightly more smear signals by subsequent Fix5-29 PCR but repeated cloning of identical sequences became more frequent at lower target DNA concentrations, thus reducing the overall efficiency of cloning representative DNA fragments. Forty-five cycles was found to be optimum for Fix5-29 PCR. Fewer than 40 cycles generated very weak smear signals at the femtogram target DNA level and reduced cloning efficiency. The cycle numbers for both reactions were determined for target DNA at lower than the 1 pg level. When sample DNA amounts are relatively higher, fewer cycles may be needed to amplify a sufficient amount of DNA for cloning.

The size of Fix5-29 primer PCR products ranges between 0.3 and 3 kb, with the majority in the 0.3–0.5 kb range gradually fading out after reaching 2 kb. To facilitate the cloning of large DNA molecules, Fix5-29 PCR products between 0.5 and 2 kb were retrieved from the gel, digested with *Pst*I, and ligated with the vector. A large number of white *E. coli* colonies (10^3 per transformation) were produced, while the number of blue colonies stayed at the background level, which suggests that most of the white colonies harbor recombinant plasmids as opposed to background white colonies. To select recombinant plasmids containing large inserts, we performed PCR with M13 forward and reverse primers flanking the pUC19 multiple cloning sites. We randomly examined 19 white colonies derived from each target DNA sample by PCR. Figure 2 shows the PCR products from white *E. coli* colonies that were derived from cloning procedures of 100, 10, and 1 pg original *M. genitalium* DNA samples. Similar gel electrophoresis results were obtained with other samples (data not shown). Each PCR band represents an insert amplified from the recombinant plasmid in a white *E. coli* colony. The average size of the inserts appeared to decrease at lower amounts of original DNA sample.

For each amount of DNA, we sequenced up to five PCR products of relatively large size with M13 forward primer. The resulting sequences were aligned with the GenBank® Database (National Center for Biotechnol-

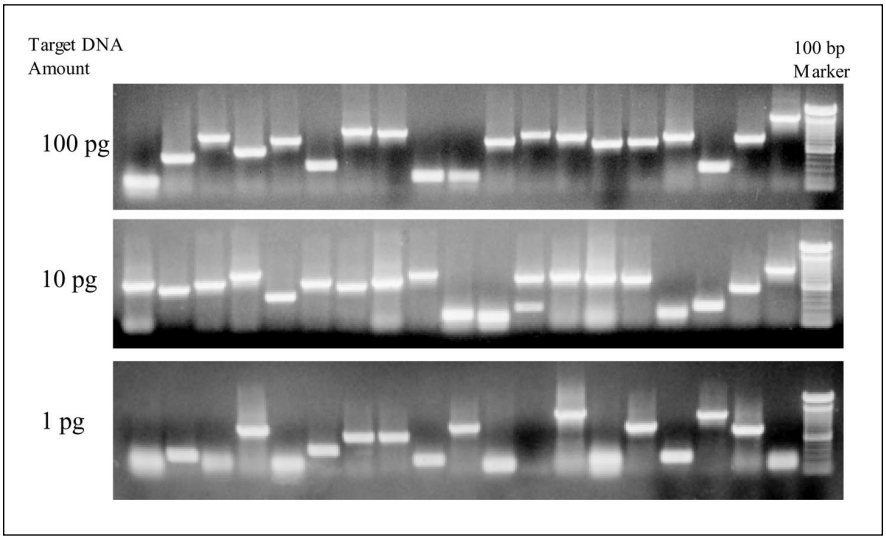

Figure 2. M13 forward and reverse PCR products with *Escherichia coli* colonies harboring recombinant plasmids. Ran5-29/Fix5-29 PCR products with target *M. genitalium* DNA samples of 100, 10, and 1 pg were cloned in pUC19 and transformed into *E. coli*. PCRs were carried out using M13 forward and reverse primers with white *E. coli* colonies. The rightmost lanes are 100-bp DNA markers.

ogy Information, National Institutes of Health, Bethesda, MD, USA) for identification. Table 1 shows the summary of the GenBank Basic Local Alignment Search Tool (BLAST®) results. At target DNA amounts of 100 and 10 pg of both human and *M. genitalium* DNA, all cloned sequences were from their respective origins. The sequences from the 1 pg human DNA sample included two groups of same human sequences, which suggests that Fix5-29 PCR products were produced from a limited variety of Ran5-29-derived templates, an expected result of low original template amounts. When we were working on 100 or 10 fg amounts of human DNA, PCR products of significant size became difficult to obtain from white *E. coli* colonies (Table 1). Human sequences were still cloned from the 100-fg DNA sample, but sequences of nonhuman origins began to appear, suggesting that our PCR system has a 100-fg nonspecific background DNA level. Only two cloned sequences could be obtained from the 19 screened *E. coli* colonies that had originated from the 10 fg target human DNA sample; however, none were of human origin (Table 1). The cloning sensitivity for *M. genitalium* DNA samples, in terms of target DNA amount, is similar to that of human DNA (Table 1).

The N7TGGC sequence introduced by PCR using the Ran5-29 primer should remain in the recombinant plasmids after the subsequent cloning procedure of the Fix5-29 PCR products. Theoretically, N7TGGC is only removed during the cloning of the PCR products that contain an internal *Pst*I site. We found most of the cloned human sequences lacked the N7TGGC sequence, whereas most of the cloned mycoplasma sequences retained the sequence marker (Table 1). We verified the sequences flanking the cloned PCR products that lacked the N7TGGC marker at the end(s) against the GenBank database. Over 50% of the cloned human sequences lacking the N7TGGC marker were flanked by *Pst*I sites in the original sequence deposited in the GenBank database, while the same situation only happened in less than 20% of the cloned *M. genitalium* sequences lacking the N7TGGC marker. This partially explained why cloned PCR products from human DNA were generally shorter than those from *M. genitalium* DNA because the frequency of finding both the TGGC sequence and *Pst*I site (CTGCAG) in human DNA should be much higher than mycoplasma DNA, which has a much smaller genome and lower GC content.

Because we know the origin of target DNA in the current study, the cloned DNA sequences can be identified directly by comparing them to the GenBank database. However, in experiments with unknown target sequences, the data obtained by alignment to GenBank provide only the information of the DNA fragments being cloned that could have been derived from the system's background. This is especially true when the DNA amount in the sample is lower than 1 pg (see Table 1). Further examination is necessary to correlate the cloned DNA sequences with samples of interest. In this situation, our approach for working with real samples is using PCR with primers designed from the obtained DNA sequences to examine target DNA samples. A consistent amplification of expected PCR products from samples, while not from negative PCR controls, will be evidence that the obtained sequence fragment originated from the samples.

Other researchers used random hexamers tagged with specific sequences at the 5′ end for random amplification (4,5). We used random heptamers and introduced a four-nucleotide TGGC clamp at the 3′ end of the random sequence. This arrangement increases the specific annealing to 11 nucleotides and therefore greatly facilitated primer binding at higher temperatures. This strategy greatly increased the amplification sensitivity achieved by tagged random hexamers. Compared with SISPA, this approach showed an over 10-fold increase in the sensitivity of cloning very small amounts of DNA. In addition, this method has the advantage of amplifying DNA in both nature and denatured forms and provides more choices of preparation methods of target DNA, which can be very difficult when working with trace amounts.

Due to its nature of random and highly sensitive amplification, this approach should be limited to dealing with trace DNA from relatively pure sources in which DNA from the background will not overcome DNA

of interest. We believe this approach will be useful to characterize DNA in samples such as cultures of fastidious microorganisms, paraffin-embedded tissues, and samples from extreme harsh locations where the quantity and variety of living organisms are very low. It can also be used to amplify trace ancient DNA sequences in fossils in which the contamination of modern DNA sequences can be eliminated.

ACKNOWLEDGMENTS

We thank Dr. Douglas Wear and Dr. Shimin Zhang for constructive discussion and review in the preparation of this manuscript.

REFERENCES

1. **Hampson, I.N., L. Hampson, and T.M. Dexter.** 1996. Directional random oligonucleotide primed (DROP) global amplification of cDNA: its application to subtractive cDNA cloning. Nucleic Acids Res. *24*:4832-4835.
2. **Akowitz, A. and L. Manuelidis.** 1989. A novel cDNA/PCR strategy for efficient cloning of small amounts of undefined RNA. Gene *81*:295-306.
3. **Reyes, G.R. and J.P. Kim.** 1991. Sequence-independent, single-primer amplification (SISPA) of complex DNA populations. Mol. Cell. Probes *5*:473-481.
4. **Wong, K.K., L.C. Stillwell, C.A. Dockery, and J.D. Saffer.** 1996. Use of tagged random hexamer amplification (TRHA) to clone and sequence minute quantities of DNA-application to a 180 kb plasmid isolated from Sphingomonas F199. Nucleic Acids Res. *24*: 3778-3783.
5. **Peng, H.Z., P.G. Isaacson, T.C. Diss, and L.X. Pan.** 1994. Multiple PCR analyses on trace amounts of DNA extracted from fresh and paraffin wax embedded tissues after random hexamer primer PCR amplification. J. Clin. Pathol. *47*:605-608.
6. **von Eggeling, F. and H. Spielvogel.** 1995. Applications of random PCR. Cell. Mol. Biol. *41*:653-670.
7. **Mao, H. and K.S. Rosenthal.** 1999. Synthesis of radioactive single-stranded DNA probes using asymmetrical PCR and oligonucleotide random priming. BioTechniques *27*:674-678.
8. **Nie, X. and R.P. Singh.** 2001. A novel usage of random primers for multiplex RT-PCR detection of virus and viroid in aphids, leaves, and tubers. J. Virol. Methods *91*:37-49.
9. **Shao, Z., H. Zhao, L. Giver, and F.H. Arnold.** 1998. Random-priming in vitro recombination: an effective tool for directed evolution. Nucleic Acids Res. *26*:681-683.
10. **Trueba, G.A. and R.C. Johnson.** 1996. Random primed gene walking PCR: a simple procedure to retrieve nucleotide fragments adjacent to known DNA sequences. BioTechniques *21*:20.
11. **Hodgson, C.P. and R.Z. Fisk.** 1987. Hybridization probe size control: optimized 'oligolabelling'. Nucleic Acids Res. *15*:629-625.

Received 15 May 2003; accepted 1 July 2003.

Address correspondence to Shyh-Ching Lo, AFIP, Bldg. 54, Room 4091, Department of Infectious and Parasitic Diseases Pathology, 6825 16th Street, NW, Washington, DC 20306, USA. e-mail: los@afip.osd.mil