



GENOME RESEARCH

Optimization of long-distance PCR using a transposon-based model system.

L D Ohler and E A Rose

Genome Res. 1992 2: 51-59

Access the most recent version at doi:[10.1101/gr.2.1.51](https://doi.org/10.1101/gr.2.1.51)

References

This article cites 19 articles, 8 of which can be accessed free at:
<http://genome.cshlp.org/content/2/1/51.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Copyright © Cold Spring Harbor Laboratory Press

Optimization of Long-distance PCR Using a Transposon-based Model System

Lynne D. Ohler and Elise A. Rose¹

Human Genome Laboratory, Perkin-Elmer Cetus Instruments, Emeryville, California 94608

The ability to amplify routinely long PCR products (5–25 kb) with high specificity and fidelity, regardless of target template sequence or structure, would provide significant benefits to genome mapping and sequencing endeavors. Although occasional reports have described the generation of long PCR products,^(1–4) such results have been difficult to replicate and have frequently utilized probe hybridization to identify the specific product from nonspecific amplified DNA. Production of specific PCR products has generally been limited to target templates of less than 3 kb.⁽⁵⁾ To extend the effective range of standard PCR amplification, it may be necessary to utilize alternative reaction conditions and/or components, such as novel thermostable DNA polymerases or accessory proteins. We describe the use of a model system to evaluate systematically methodological changes that might enable efficient long-range PCR. Specifically, the transposon Tn5supF has been used to introduce randomly identical, known primer binding sites within separate isolates of phage clones carrying identical inserts.⁽⁶⁾ Transposon-based PCR allows us to study amplification of DNA fragments that vary in size and sequence using only a single set of primers. In the present studies, we describe conditions that enable PCR amplification of specific DNA templates ranging in size up to 9 kb. Some of the key features of our methodology include the use of recombinant *Thermus thermophilus* (rTth) DNA polymerase, the addition of gelatin to the reaction mixture, the use of wax-mediated “hot

starts”⁽⁷⁾ and, lastly, the use of auto-segment extension thermocycling. These results also provide insights into additional approaches that might further enhance our ability to perform long-distance PCR.

The versatility and power of the polymerase chain reaction (PCR) has encouraged its involvement in almost every aspect of genome mapping and sequencing.⁽⁸⁾ The application of PCR has been central in formulating both the conceptual and practical approaches on which the entire Human Genome Project is based, as evidenced by the adoption of the sequence tagged sites (STS) proposal.⁽⁹⁾ Other genomic mapping applications of PCR technology include its use in isolation of human DNA fragments present in hybrid cell lines or from microdissected or flow-sorted chromosomal regions⁽¹⁰⁾; chromosomal walking and expansion of contiguous yeast artificial chromosome (YAC), cosmid, or phage clones^(11,12); fingerprinting of DNA fragments for ordering and orienting YAC, cosmid, or phage clones⁽¹⁰⁾; fingerprinting of DNA in hybrid cell lines for identification (and to check for stability of human DNA content)⁽¹⁰⁾; mapping and analysis of expressed sequences from a particular genomic region⁽¹³⁾; and direct PCR-coupled DNA sequencing.^(14,15)

As efforts to map and sequence complex genomes continue, there is an ever-increasing need to develop streamlined methodologies that will reduce the time

and cost involved in laboratory research. In addition, focus will shift to analysis of genomic regions that have been recalcitrant to current mapping and sequencing approaches. Such regions include those that are especially rich in guanine (G) and cytosine (C), or contain complex secondary structures. Efficient amplification through these regions would facilitate analysis and characterization of many of the gaps in current genomic maps.

The ability to amplify routinely templates as large as 5–25 kb, regardless of DNA sequence or structure, would represent a major breakthrough in human genome mapping research. Such “long-range PCR” would maintain continuity, order, and orientation of DNA fragments and would provide longer templates for subsequent sequencing endeavors. Moreover, reliable amplification over long distances would decrease the number of PCR amplifications that would need to be performed to analyze any given genomic region, eliminating the need for additional primers and other reagents, and thus decreasing both the time and cost involved in genome mapping efforts. Long-range PCR would also greatly reduce the number of overlapping regions that would need to be compared, as well as the need for extensive subcloning.

Current limitations in the size of PCR products that can be generated reliably are on the order of 3–4 kb.⁽⁵⁾ Although occasional reports have appeared describing detection of large PCR products,^(1–4) little effort has been made to optimize conditions for long-distance PCR. As a consequence, such results have been difficult to replicate and frequently required probe hybridization to identify the specific product from the nonspec-

¹Present address: Advanced Center for Genetic Technology, Applied Biosystems Inc., 850 Lincoln Centre Drive, Foster City, CA 94404.

cific amplified DNA. Conditions have yet to be described that enable reliable amplification with high specificity and fidelity of long PCR products, regardless of the target template sequence.

Our long-term goal is to explore a variety of thermostable DNA polymerases under different reaction conditions that might enable (1) PCR amplification of longer products (5–25 kb) with high specificity and fidelity, and (2) PCR amplification from regions that are currently difficult to amplify (due to either complex secondary structure and/or high G + C content). Our model system employs transposons to introduce unique known primer binding sites into any targeted region of DNA, thus making the region amenable to PCR amplification and subsequent sequencing.⁽⁶⁾ Transposons are discrete DNA segments that move (or transpose) relatively randomly from one DNA site to another in the absence of apparent DNA homology. Transposon-based PCR represents a potential method for inserting PCR primer binding sites into regions that are difficult to analyze and for which no known sequence information is available.

In the current studies, the transposon *Tn5supF* has been used to introduce PCR primer binding sites at random locations within separate isolates of phage λ 138 clones carrying identical *Escherichia coli* inserts.^(4,6) PCR-mediated mapping of transposon insertions has been carried out using primers specific for *Tn5supF* inside and outside in relation to the left and right λ cloning sites. Using this system, only two sets of primers are required to study amplification of many DNA fragments that vary in size and sequence. Our use of *Tn5supF* transposon-based PCR as a model system has enabled us to examine a variety of new reaction components and conditions, with primer binding sites spaced at increasing distances, without having to change PCR primers. This approach has already enabled us to define conditions that extend the range of PCR to 9 kb.

MATERIALS AND METHODS

DNA Templates

DNA templates used in these studies were λ phage 138 clones carrying the *E. coli lacZ* gene with a small amount of surrounding *E. coli* DNA. These clones

were kindly provided by D. Berg (Washington University, St. Louis). A single *Tn5supF* transposon was randomly inserted at a unique location within each cloned fragment.^(4,6) Phage templates designated 4, 15, S lac, 3 lac, 2, 14, and 9 carried transposon insertions at the sites indicated in Figure 1. The bacterial strains MC1061 or LE392 were used interchangeably for plating of λ phage. Plaque isolates were dissolved in 50 μ l of SM buffer⁽¹⁶⁾ and stored at 4°C. Aliquots of 5 μ l were used for PCR amplifications.

PCR Primers

PCR was performed using primers that are complementary to unique sequences

near each transposon end and with primers specific for the two ends of the vector. The primers and their sequences are as follows: λ left arm cloning site (L), 5'-ATAGAGTCTTGACAGACAAACTG-C-3'; λ right arm cloning site (R), 5'-GCCTAACGATCATATACATGG-3'; *Tn5supF* inside end (I), 5'-TAGGATCCCAGATC-TGATC-3'; and *Tn5supF* outside end (O), 5'-TAGGATCCCCTACTTGTGTA-3'. Primers I and O were designed by Phadnis et al.⁽⁶⁾ Oligonucleotides L and R were designed using annealing thermodynamics software developed by D. Birch, Q. Chou, and W. Bloch (Cetus Corp). Use of primers with similar melting temperature (T_m) proved to be critical in obtaining specific amplification of desired long PCR products.

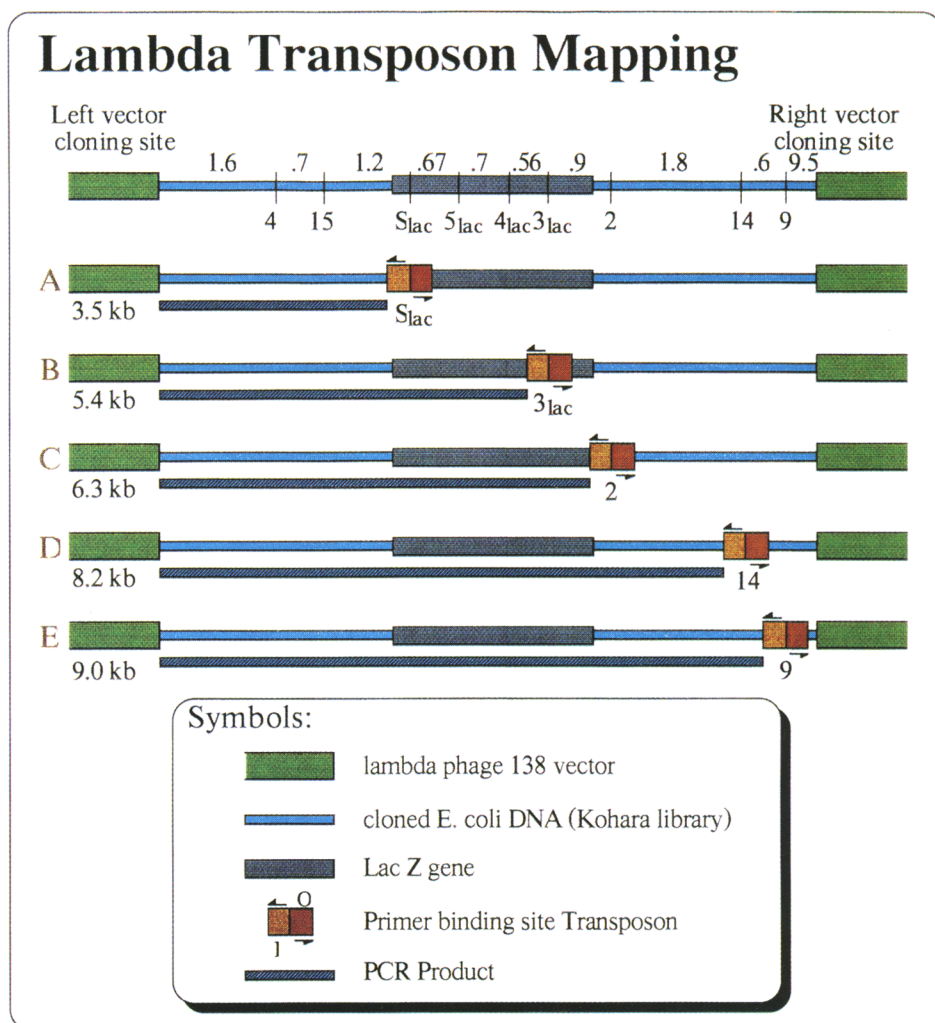


FIGURE 1 Mapping of transposon insertions. The model system used in these studies employs introduction of the transposon *Tn5supF* into random locations of individual phage λ 138 clones. Each phage clone carries the *E. coli lacZ* gene with a small amount of surrounding DNA.⁽⁶⁾ Mapping of transposon insertions is carried out using primers specific for *Tn5supF* inside (I) and outside (O) ends and for λ DNA adjacent to the left (L) and right (R) cloning sites.⁽⁶⁾

PCR Amplification Reaction Components

PCR amplification reactions were performed in 50- μ l total volumes. Each reaction mixture contained 5 μ l of sample DNA template 250 μ M each dNTP (Perkin-Elmer Cetus), and 30 pmoles each of two primers (either L or R, along with either I or O). Primers were added to the reaction mixture as either unlabeled oligomers, or labeled with JOE or FAM fluorescent dyes (Applied Biosystems, Inc.). A variety of DNA polymerases were employed in different reactions, along with the appropriate buffers, with each enzyme tested at several concentrations. These enzymes included: *Thermatoga maritima* (*Tma*) DNA polymerase (kindly provided by D. Gelfand, Cetus Corp.), 1–2 units/50 μ l, in 1 \times PCR buffer (consisting of 50 mM KCl, 10 mM Tris-HCl, pH 8.3) supplemented with 2.5 mM MgCl₂; AmpliTaq DNA Polymerase (recombinant *Taq* DNA polymerase), 1.25–15 units/50 μ l, in 1 \times GeneAmp PCR Buffer (50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM MgCl₂, 0.001% gelatin) supplied by the manufacturer (Perkin-Elmer Cetus); recombinant *Thermus thermophilus* (*rTth*) DNA polymerase, 0.625–7.5 units/50 μ l, in 1 \times *rTth* PCR buffer (5% glycerol, 10 mM Tris-HCl, pH 8.3, 100 mM KCl, 0.75 mM EGTA, 0.05% Tween 20) supplied by the manufacturer (Perkin-Elmer Cetus), supplemented with 2.5 mM MgCl₂; Hot Tub DNA Polymerase, 0.4–5.0 units/50 μ l, in the 1 \times primer extension buffer (50 mM Tris HCl, pH 9.5, 1.5 mM MgCl₂, 20 mM (NH₄)₂SO₄) supplied by the manufacturer (Amersham); VENT DNA Polymerase (*Thermococcus litoralis* DNA polymerase), 0.25–1.25 units/50 μ l, in the 1 \times primer extension buffer (10 mM KCl, 10 mM (NH₄)₂SO₄, 20 mM Tris-HCl, pH 8.8, 2 mM MgSO₄, 0.1% Triton X-100) supplied by the manufacturer (New England Biolabs); AmpliTaq DNA Polymerase, Exo⁻ Mutein (recombinant mutant *Taq* DNA polymerase containing 1 amino acid substitution), 0.5–2.0 units/50 μ l, in 1 \times GeneAmp PCR Buffer II with MgCl₂ as described above for *Tma* (kindly supplied by R. Abramson and D. Gelfand, Cetus Corp.).

In addition to examining a variety of DNA polymerases and polymerase concentrations, we also systematically evaluated the effects of a range of concentrations of reaction components that might act to “stabilize” DNA polymerases.

These reagents included Tween 20 (0.001–1.0%) and gelatin (0.005–0.05%).

Hot Start Technique

The specificity of PCR amplification can be greatly enhanced by minimizing primer extension at misprimed target sites and reducing the rate at which primer oligomerization occurs at non-stringent temperatures. Nonspecific PCR products are typically generated prior to PCR, and during the early stages of thermal cycling. Production of the desired specific PCR product can be favored by initially keeping at least one critical reaction component separate from the rest of the reaction mixture until the temperature has exceeded approximately 80°C, and then adding the missing component(s) so that amplification begins at an elevated temperature. This general approach to improving the specificity of PCR amplification is referred to as the “hot start” technique.^(7,17,18) In the present studies, we used 12 mg AmpliWax PCR Gems (kindly provided by W. Bloch, Cetus Corp.) to facilitate automated hot starts. This product is a thermolabile material that is added to an initial reaction mixture containing primers, dNTPs, and buffer. The tubes are capped, placed in the thermal cycler, and heated to 80°C for 5 min. At this elevated temperature, the PCR Gems melt to form a liquid layer on top of each reaction mixture. The temperature is then reduced to 25°C for 1 min, during which time the PCR Gems resolidify and form a barrier covering the PCR mixtures. The reaction tubes are then opened, and an upper reaction mixture containing the DNA polymerase and sample (and in some cases, additional experimental reaction components) are layered on top of the solid wax barrier. The tubes are recapped and subjected to PCR as described below. During the initial cycle of PCR, the barrier again melts, allowing the thermostable enzyme and sample DNA to mix with the rest of the reaction components. Following this hot start, PCR proceeds normally with the AmpliWax PCR Gems subsequently acting as a vapor barrier above the reaction mixture.

PCR Amplification Thermal Cycling Conditions

DNA thermal cycling was initially performed using both the DNA Thermal Cy-

cler 480 (Perkin-Elmer Cetus) and the GeneAmp PCR System 9600 (Perkin-Elmer Cetus). Optimization of reaction conditions, however, was carried out using the System 9600 exclusively. For these experiments, PCR amplification reaction mixtures were prepared in thin-walled MicroAmp reaction tubes (Perkin-Elmer Cetus). A variety of cycling conditions were investigated. Denaturation temperatures of 93°C and 95°C for 10 sec were used. In addition, annealing temperatures ranging from 58°C and 74°C, and annealing time ranging from 15 sec to 30 sec, were tested for each set of primers. The “auto segment extension” program was varied by altering the number of standard cycles (10–15) and automatically extended cycles (15–20), as well as the amount of extension time added (15–30 sec) at each automatically extended cycle. Optimal results were obtained when samples were subjected to an initial 10 cycles each consisting of: 95°C for 10 sec, 58°C for 30 sec, and 72°C for 3 min. This was followed by 20 cycles, each consisting of: 95° for 10 sec, 58°C for 30 sec, and 72°C for 3 min plus a 30-sec automatic increment every cycle.

Samples were stored at 4°C following PCR, prior to analysis.

Gel Electrophoresis and Transposon Mapping

PCR product size was determined by gel electrophoresis, ethidium bromide staining, and either ultraviolet or fluorescent detection. On occasion, fluorescent PCR products were also generated by incorporation of fluorescent dye-tagged primers into PCR product during the amplification reaction. Fluorescent PCR products were analyzed using the Gene Scanner 362 Fluorescent Fragment Analyzer (Applied Biosystems, Inc.). This system was chosen because it allowed us to analyze very small portions of our PCR reaction mixtures, leaving the remainder available for other experiments such as long-distance DNA sequencing. Typically, aliquots of 1–5 μ l of each 50 μ l PCR mixture, along with 2 \times gel loading buffer (Applied Biosystems, Inc.), were loaded onto a 0.8% agarose gel in the Gene Scanner. The samples were subjected to electrophoresis for 6 hr in 1 \times TBE buffer, at ambient temperature. PCR products were detected based on either the inherent fluorescence of ethidium

bromide, or the fluorescence generated by the JOE or FAM labeled primers. When detection relied on ethidium fluorescence, 0.05 μ mole ethidium bromide was included in both the buffer and the gel, and a 1-kb DNA ladder (BRL/Life Technologies, Inc.) was used as size markers. When JOE or FAM dye-labeled primers were used, GENESCAN-30000 size markers labeled with ROX dye (Applied Biosystems, Inc.) were often run in the same lanes as the samples, providing lane-specific sizing of PCR products.

RESULTS

Transposon Insertions

TN5supF insertions into individual phage clones, designated 4, 15, S lac, 3 lac, 2, 14, and 9, were mapped using primers complementary to unique DNA sequences adjacent to each transposon end and with primers specific to one of the two ends of the vector (Fig. 1). This system was used in all investigations of amplification reaction components and conditions.

Primer Design

Our initial studies involved use of primers previously described by Krishnan et al.⁽⁴⁾ Use of these primers consistently resulted in generation of significant nonspecific products, and made amplification of the desired PCR product difficult to replicate (Fig. 2). Upon further investigation, we determined that the primers in each of these pairs had melting points (T_m) that differed by as much as 10°C. Therefore, new I and O primers were designed, as described in the Materials and Methods section above, which had melting points identical to the L and R primers, respectively (i.e., 73.7°C in 50 mM KCl, and 69.1°C in 50 mM KCl, respectively). These new primers were used in all subsequent efforts to optimize long-distance PCR in this model system. We note that differences in the T_m of primer pairs are not of consequence in routine "short distance" PCR, but that nonspecific PCR products do occur when the extension times are increased.

Evaluation of DNA Polymerases

Ranges of concentrations for a variety of thermostable DNA polymerases were in-

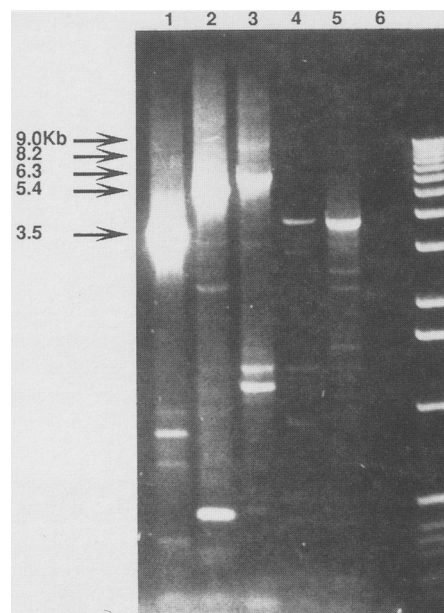


FIGURE 2 Effects of primer design on specificity of PCR. (Lanes 1–5) PCR products generated following amplification of the 3.5-, 5.4-, 6.3-, 8.2-, and 9.0-kb targets, respectively, using the L and I primers described in Krishnan et al.⁽⁴⁾ The T_m of these primers differed by approximately 10°C. (Lane 6) Negative control obtained by performing PCR on a reaction mixture containing no DNA. Aliquots of 20 μ l (out of the 50 μ l PCR reaction) were loaded onto a 0.8% agarose gel. Note the presence of many nonspecific bands and background smearing in all lanes, as was seen in the original paper,⁽⁴⁾ and the fact that the 8.2- and 9.0-kb bands are barely visible.

vestigated to define conditions resulting in consistent amplification of long PCR products. The specific enzymes evaluated were chosen due to their different inherent exonuclease activities, under the assumption that such activities might either enhance or limit a polymerase's ability to extend a primer over long distances. The enzymes used, and their respective exonuclease activities,

are described in Table 1. All of these enzymes generally enable PCR amplification of fragments ranging up to 3.5 kb. Larger PCR products (3.5–6.0 kb) were occasionally generated with AmpliTaq DNA polymerase, Hot Tub DNA polymerase, the AmpliTaq Exo⁻ mutein, and *rTth* DNA polymerase (data not shown). Of these enzymes, *rTth* DNA polymerase most consistently yielded PCR products larger than 3.5 kb. Consequently, subsequent approaches to extending the range of PCR were investigated using this enzyme. We did not attempt to optimize PCR further with the remaining collection of enzymes. We believe, however, that several of these enzymes may be useful in long-range PCR, either alone or in combination with another polymerase, if reaction conditions are optimized.

Modification of Other Reaction Components

We investigated possible effects of reagents that may act to "stabilize" DNA polymerases.⁽¹⁹⁾ *rTth* DNA polymerase has a half-life of approximately 20 min at 95°C.⁽²⁰⁾ Given the long incubation times required at elevated temperatures for *rTth* DNA polymerase to extend a primer sufficiently to obtain a 10-kb PCR product, we were concerned that a critical level of enzyme activity might be compromised during a 30-cycle PCR. We hoped that the addition of either the detergent Tween 20 or gelatin, both known polymerase stabilizers,⁽¹⁹⁾ might help to maintain optimal polymerase activity, thereby increasing the likelihood that the polymerase would be able to extend over longer distances. No effect was seen with the range of Tween 20 used (0.001–1.0%). The addition of gelatin at 0.01%, however, had a significant effect that enabled routine generation of 9.0-kb PCR

TABLE 1 Exonuclease Activities of DNA Polymerases

| DNA polymerase | 5' → 3' Exonuclease | 3' → 5' Exonuclease | References |
|----------------------------------|------------------------|------------------------|------------|
| <i>Tma</i> | + | + | 27 |
| AmpliTaq | + | – | 28 |
| <i>rTth</i> | + | – | 20 |
| Hot Tub | + | – | 29 |
| VENT | – | + | 30 |
| AmpliTaq Exo ⁻ Mutein | – | – | 26 |

+ indicates that enzyme possesses activity. – indicates that enzyme lacks activity.

products (Figs. 3 and 4). Gelatin was therefore used in all subsequent PCR experiments.

Although we were able to amplify larger PCR products, we occasionally observed nonspecific PCR products and generation of primer-dimer artifacts. This problem was alleviated through the use of wax-mediated hot starts. AmpliWax PCR Gems proved to be an effective and convenient tool for implementing this technique.⁽⁷⁾ In addition to providing increased specificity in PCR, the use of wax-mediated hot starts also increased yield in PCR amplifications. Moreover, utilization of this methodology provided greater flexibility in experimental design by permitting specific PCR to occur at temperatures that might otherwise have resulted in the generation of nonspecific products. The time spent on temperature optimization was thus minimized.

Cycling Conditions

Optimal conditions for long-distance PCR were designed for use on the System

9600. Initial experiments using the DNA Thermal Cycler 480 yielded inconsistent results. No effort was made, however, to optimize PCR conditions for use on this instrument. Other types of DNA Thermal Cyclers were not used due to lack of availability in our laboratory.

One advantage of using the System 9600 was the ability to program the instrument to increase incubation times automatically in the later cycles of a reaction. DNA polymerases extend primers discontinuously through a succession of reactions, moving on to and off from the DNA template. As PCR product begins to accumulate in the later cycles, the ability to extend a significant proportion of primers over a long distance might be limited by the relative decrease in enzyme molecules per template. The "auto segment extension" software feature of the System 9600 enabled us to increase extension time automatically by 30 sec every cycle following the tenth cycle. We speculated that prolonging the extension time in the later cycles might provide sufficient time for the enzyme to

further extend primers through new template DNAs. Whether or not this rationale is correct, we found that use of autosegment extension resulted in more consistent generation of long PCR products.

Optimal Conditions for Long-range PCR

Fifty-microliter (*final volume*) reaction mixtures were prepared in MicroAmp tubes with the following components (at the specified *final* concentrations) added to the tubes prior to melting the AmpliWax PCR Gems: 30 pmoles primer I or O; 30 pmoles primer L or R; 2.5 mM MgCl₂; 1.0 mM total dNTPs; 25 mM Tris HCl, pH 8.9; one AmpliWax PCR Gem; in an initial volume of 13.5 μ l. After melting the AmpliWax PCR Gem and allowing the thermolabile barrier to resolidify above the lower reaction mixture, the following components were added to form an upper layer in the reaction tubes: 1 \times *rTth* PCR buffer; 2.5 units *rTth* DNA polymerase; 0.01% gelatin; 5.0 μ l of ph-

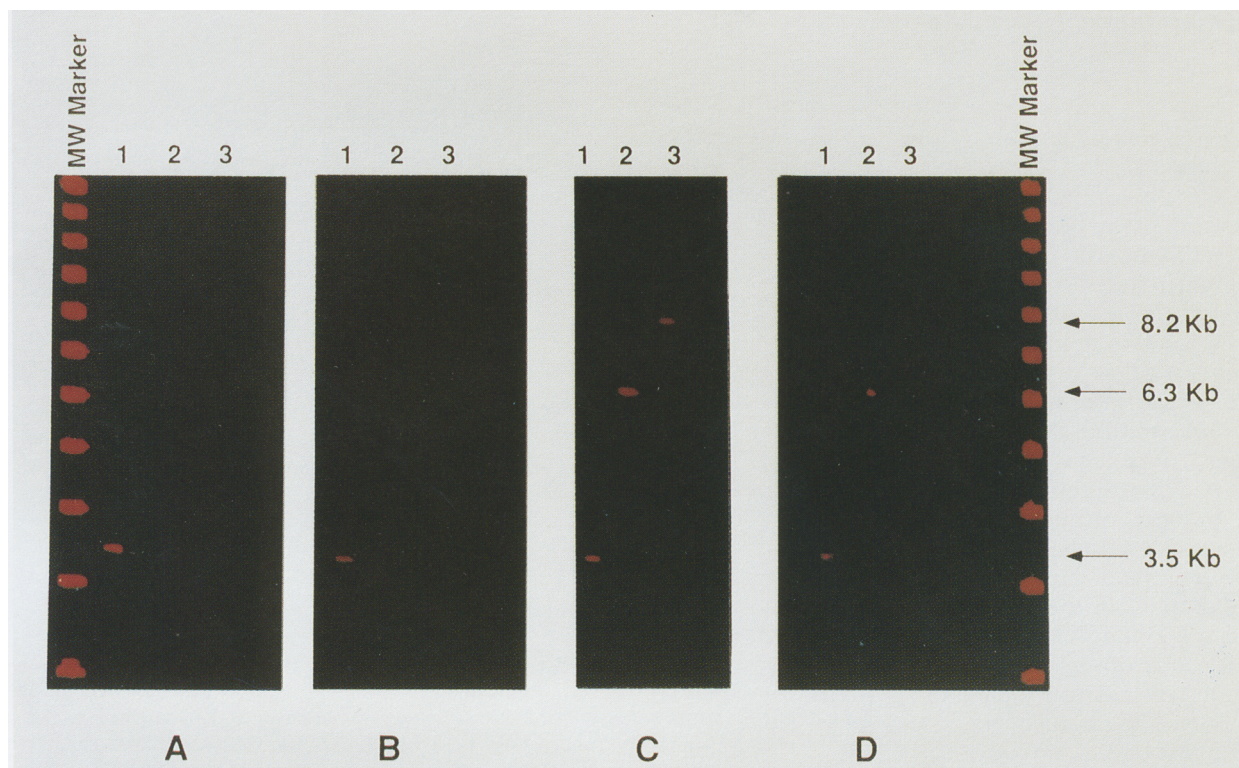


FIGURE 3 Effects of adding gelatin to reaction mixtures. Utilizing the ABI Gene Scanner 362 Fluorescent Fragment Analyzer, a profile of PCR fragments ranging from 3.5 kb to 8.2 kb is shown in the photograph. (A, B, C, and D) PCR products generated using reaction mixtures which contain 0%, 0.005%, 0.01%, and 0.05% gelatin, respectively. (Lanes 1–3) PCR products of 3.5, 6.3, and 8.2 kb and correspond to the λ Tn5*supF* isolates A, C, and D depicted in Fig. 1, respectively. Only the 3.5-kb PCR product was detectable when 0% or 0.005% gelatin was added to each sample. A concentration of 0.01% gelatin yielded the highest molecular weight products (C). The higher concentration of 0.05% gelatin appeared to inhibit the PCR.

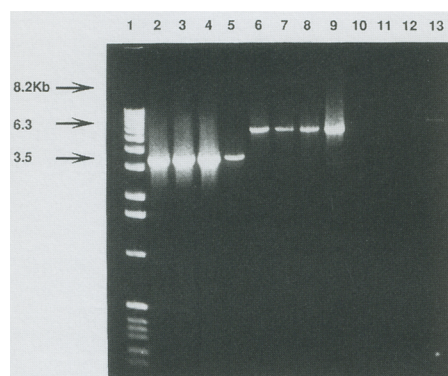


FIGURE 4 Effects of adding gelatin to reaction mixtures, as visualized by UV detection of ethidium bromide-stained DNA. Aliquots of 20 μ l of each PCR reaction were loaded onto a 0.8% agarose gel. (Lane 1) Molecular weight size marker; (lanes 2–5) PCR products generated from amplification of a 3.5-kb target; (lanes 6–9) PCR products generated from amplification of a 6.3-kb target; (lanes 10–13) PCR products generated from amplification of a 8.2-kb target. Reaction products in lanes 2, 6, and 10 were generated using 0% gelatin; 0.001% gelatin was used to produce the PCR products in lanes 3, 7, and 11; 0.005% gelatin was used to obtain the PCR products in lanes 4, 8, and 12; and 0.01% gelatin was used for the reaction products run in lanes 5, 9, and 13. Once again, optimal long-distance PCR results were obtained when reaction mixtures contained 0.01% gelatin.

age plaque; in a volume of 36.5 μ l. Samples were then placed in the GeneAmp PCR System 9600, and amplified using cycling conditions as described in the Materials and Methods section above.

In the model system tested, specific PCR products up to 9 kb in size were obtained when reaction mixtures were formulated and used as described above (Fig. 5). Following optimization of reaction conditions, amplification of specific PCR products as large as 8.2 kb in size were obtained faithfully (Fig. 6). The 9-kb PCR product, although generally obtained, was not always able to be detected. This might indicate a low yield of the desired product, or that further optimization is still required. We have yet to define PCR conditions enabling reliable amplification targets larger than 9 kb.

Amounts of PCR products generated varied with size of the amplified DNA fragment. Based on analysis of ethidium-stained gels, for example, we estimate that amplification of the 3.5- and 5.4-kb targets yielded approximately 1–5 μ g of DNA, the 6.0-kb target yielded 1–2 μ g, the 8.2-kb fragment yielded 200–500 ng,

and the 9.0-kb product yielded 50–150 ng.

DISCUSSION

The long-term objective of our research is to evaluate, develop, and improve upon applications of the PCR for genomic mapping and sequencing.^(8,21–23) These efforts involve the systematic evaluation of a variety of thermostable DNA polymerases, PCR amplification reaction components, and reaction conditions. In the course of this research we hope to develop methods that will enable one to (1) amplify large PCR products reliably with high specificity and fidelity, and (2) amplify regions that are currently recalcitrant to PCR, due to complex secondary structure and/or high G + C content.

The ability to perform long-range PCR (i.e., specific amplification of targets in the 5- to 25-kb range), irrespective of the DNA sequence or structure of the target template, will help close gaps in existing maps, potentially eliminate the need for subcloning of large DNA fragments, and maintain continuity of order and orientation of DNA fragments that are being mapped and sequenced. Long-range PCR will thus streamline genome mapping efforts as well as provide a source of PCR products to serve as templates for long-distance DNA sequencing (E.A. Rose, in prep.).

Although other investigators have reported detection of large PCR products,^(1–4) little effort has been made to optimize conditions specific for long-distance PCR. As a consequence, early at-

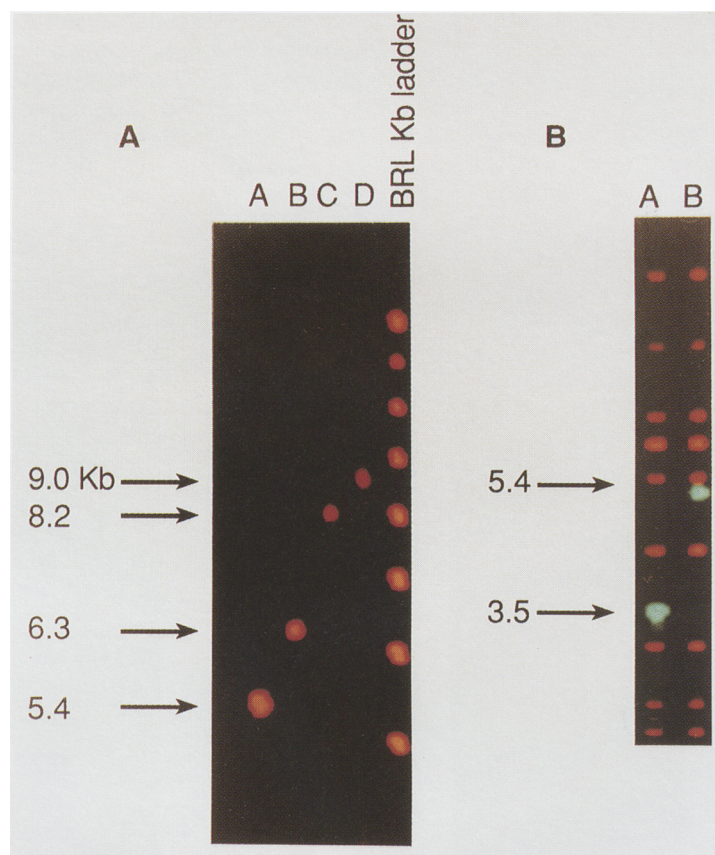


FIGURE 5 Effect of optimal PCR amplification conditions. (A) Utilizing the ABI Gene Scanner, a profile of PCR fragments ranging from 3.5 kb to 9.0 kb is shown in the photograph to the left. (Lanes A–D) Correspond to the λ 138::Tn*SsupF* isolates B–E, as depicted in Fig. 1. A total of 1–5 μ l of each 50 μ l PCR mixture and ABI 2 \times gel loading buffer were loaded onto a 0.8% agarose gel. The gel was run for 6 hr in 1 \times TBE buffer, at ambient temperature, with 0.05 μ mole ethidium bromide included in both buffer and gel. PCR products were detected using the inherent fluorescence of ethidium bromide. (B) Detection of PCR products using primers conjugated with JOE (Applied Biosystems). GENESCAN-30000 labeled with ROX (Applied Biosystems) is run in the same lane as the sample to serve as a size marker as well as a quantitative marker. The same reaction and electrophoretic conditions were used in both systems. (Lanes A and B) Correspond to λ 138::Tn*SsupF* isolates A and B shown in Fig. 1.

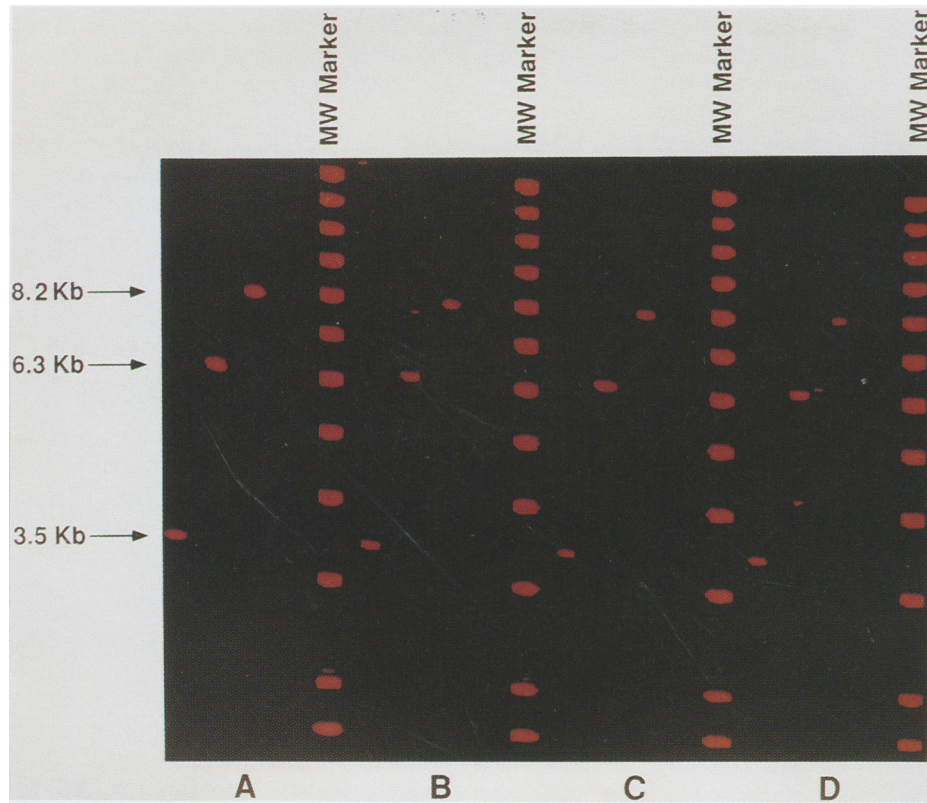


FIGURE 6 Reproducibility of long-distance PCR. (A–D) Replicate series of amplification reactions performed using the optimal conditions described in the Results section. Each panel contains three lanes in which 3.5-, 6.3-, and 8.2-kb targets, respectively, have been amplified. As can be seen, once reaction conditions were optimized, generation of specific large PCR products was highly reproducible.

tempts have been difficult to replicate and frequently required probe hybridization to identify the specific product from the nonspecific amplified DNA. Krishnan et al.,⁽⁴⁾ using the same transposon-based system as we use in this paper, report the occasional detection of high-molecular-weight PCR products. Although the authors speculate that their inability to obtain large PCR products consistently is due to the lack of availability of “mapping-grade enzymes,” examination of their methods and the resulting data indicate that PCR conditions were not optimized. Certain hallmarks of nonoptimized PCR were evident, including excessive nonspecific product generation, background “smearing,” and inconsistent ability to produce desired targets. The use of primers in the same PCR reaction that differ in melting points by as much as 10°C appeared to enhance the generation of nonspecific products. Moreover, use of denaturation incubation times of only 20 sec in a DNA Thermal Cycler is unlikely to result in adequate strand separation

because the instrument is not specified to reach the target denaturation temperature unless a 60-sec incubation is employed.

Our present studies demonstrate the utility of transposon-based PCR as a generic method for both optimizing long-distance PCR as well as for analyzing any targeted segment of genomic DNA into which transposons can be inserted. In addition, we provide evidence that specific large PCR products can be generated reliably if reaction conditions are optimized. The advantages of transposon-based PCR to genome mapping and sequencing are numerous. This methodology requires use of a limited number of primers regardless of the target template being amplified and/or sequenced. The number of variables that must be optimized is thereby limited, saving both time and cost. This was clearly an asset in the present studies, as efforts to extend the range of routine PCR amplification requires investigation of many parameters independent of primer sequences. The need to design optimized

primer sequences for each larger fragment we tried to amplify would have significantly impeded our progress, as well as added an extra degree of complexity to the problem at hand. Perhaps most importantly, however, the use of transposon-based PCR enables analysis of any target segment of DNA, without prior sequence knowledge, and makes these targets accessible for both amplification and sequencing via a single system.⁽²⁴⁾

We have evaluated a variety of thermostable DNA polymerases, PCR amplification reaction components, and PCR conditions for their ability to generate long PCR products. The enzymes chosen for investigation differ in a number of potentially important activities including extension rate, processivity, fidelity, thermostability, thermal activity profile, 3′ → 5′ exonuclease activity (proofreading activity), and 5′ → 3′ exonuclease activity (nick translation activity), many of which might be expected to affect the ability to amplify over long segments of DNA. In our hands, under the conditions described above, recombinant *Tth* DNA polymerase provided the most consistent performance in long-range PCR. The other enzymes evaluated may also be capable of generating long PCR products under certain conditions,^(1–4) but our initial results indicated that efforts would best be directed toward further optimization using *rTth* DNA polymerase.

Recombinant *Tth* DNA polymerase has a half-life of 20 min at 95°C, optimal DNA polymerase activity in the same temperature range at which stringent primer annealing occurs (65–75°C), 5′ → 3′ exonuclease activity, and no detectable 3′ → 5′ exonuclease activity.^(20,25) The enzyme’s processivity is 30–40 nucleotides in 100 mM KCl, and its extension rate is approximately 60 nucleotides/sec in 100 mM KCl.⁽²⁶⁾ Fidelity of this enzyme has yet to be determined under any assay conditions, but based on sequence homology and the lack of 3′ → 5′ exonuclease activity one would expect fidelity to be similar to that described for *Taq* DNA polymerase.⁽²⁷⁾

We speculated that optimization of long-range PCR would benefit from the use of approaches that maximize the specificity of PCR. Such methods should ensure that all available primers, nucleotides, and enzyme molecules would be used to generate the desired product. Wax-mediated hot starts enhance PCR

specificity by eliminating production of nonspecific products due to pre-PCR mispriming and primer-oligomerization (primer-dimer) during the initial steps of PCR.⁽⁷⁾ We find that uniformity and yield of PCR also improve using this approach. It has been speculated that the uniformity effect is most likely due to providing a vapor barrier of consistent mass in each reaction tube, and that the enhanced yield results from focusing all reagents on production of the specific desired product.⁽⁷⁾ By permitting specific PCR to occur at temperatures that would otherwise result in the formation of nonspecific products, wax-mediated hot starts provide the additional benefit of decreasing the time required for PCR optimization. The importance of this last point cannot be overstated when one considers the number of PCR optimization experiments that will be required to develop PCR assays for thousands of STSs.

We also speculated that long-distance PCR would be facilitated by the use of reaction components that could possibly stabilize the enzyme, act to maintain enzyme contact with the target template, and/or simply prevent the enzyme from sticking to the walls of the reaction tube. Although we cannot be sure of the mechanism underlying the effect, we found that the addition of 0.01% gelatin to the reaction mixture improved long-distance PCR.

As PCR product begins to accumulate during cycling, the relative concentration of DNA polymerase decreases (i.e., there are the same number of enzyme molecules competing for an ever-increasing number of templates). Given the processivity of *rTth* DNA polymerase (i.e., the fact that the enzyme drops off a template after extending approximately 30–40 nucleotides, and then attaches to another or the same template), and the long distances over which we were attempting to amplify, we speculated that increasing the extension time in each of the later PCR cycles would increase the likelihood that we would successfully synthesize long PCR products. We were able to program the thermal cycler to perform this function by using the “auto segment extension” feature. As noted above, we believe that use of this feature was one of the key reaction conditions enabling consistent generation of long PCR products. The other important feature provided by this system is rapid heat

transfer. This limited the amount of time that the samples were exposed to high temperatures (minimizing loss of DNA polymerase activity), as well as the time that samples were exposed to low transition temperatures (at which point nonspecific priming might occur).

The use of the ABI 362 Fluorescent Fragment Analyzer enabled rapid and sensitive detection and sizing of PCR products. A major asset of this system was the ability to use internal size standards. Another advantage of using this system is that it allowed us to analyze only 1–5 μ l of a 50- μ l PCR sample mixture, leaving the remaining PCR product available for other post-PCR procedures such as hybridization, cloning, or DNA sequencing.

Overall, we believe our ability to optimize PCR conditions for generation of large PCR products relied on the following factors. First, the use of recombinant *Tth* DNA polymerase, an enzyme having optimal activity in the temperature range at which specific primer annealing occurs, an extension rate (approximately 60 nucleotides per second) capable of extending a primer more than 10 kb in 3 min, and adequate thermal stability to remain active over 30 or more cycles during which the enzyme is transiently exposed to temperatures as high as 95°C. Second, the fact that *rTth* DNA polymerase lacks 3' \rightarrow 5' exonuclease activity, thus minimizing the likelihood that primers and single-stranded template are destroyed during PCR. Third, the use of gelatin to enhance either enzyme stability or accessibility of enzyme to DNA. Fourth, the use of primers designed to melt at the same temperature. Fifth, the enhanced specificity obtained using wax-mediated hot starts. Sixth, the ability to provide sufficient extension time, and longer extension times with increasing template concentrations, by using the auto segment extension program.

Although we have already demonstrated reliable amplification of specific DNA segments ranging up to 9 kb in this system, we are continuing to examine a variety of reaction components and conditions that we believe will enable routine amplification of even larger DNA targets. The results presented here represent a preliminary analysis of many variables simultaneously in a very complex system, and is by no means exhaustive. Therefore, we find these initial results very encouraging, and believe that they

provide indications of what to try next to extend further the limits of PCR. Factors that we plan to investigate further include cycling parameters, single-stranded DNA binding protein-mediated hot starts, different types and concentrations of salt and metal ions (e.g., KCl, NaCl, Mg²⁺, Mn²⁺, etc.), additional thermostable DNA polymerases (including mutant enzymes with alterations to specific active sites), thermostable replication enzymes (as opposed to repair enzymes, which may be represented by the majority of thermostable DNA polymerases currently available), and the combination of such thermostable replication enzymes and accessory proteins (e.g., thermostable helicases and single-stranded DNA displacement proteins). In addition, we plan to investigate a variety of additional human DNA inserts of varying sequence nature to determine whether reaction components and conditions can be identified that are generally applicable to PCR amplification of any long DNA target.

There clearly exist complex conditions enabling specific replication of long stretches of DNA in vivo. For routine PCR amplification of large target templates, this process must be simulated in vitro. The first step toward defining the conditions under which such replication can be created is to find an appropriate model for testing new PCR reaction conditions and components. We believe that the transposon-based PCR system described here is such a model. Our next step will be to find ways to amplify reproducibly 10- to 25-kb segments of DNA, regardless of sequence nature, with high specificity and fidelity. Using the types of reagents, model system, and instrumentation described here, we are in a position to evaluate systematically a wide variety of parameters in search of conditions that will enable highly specific and highly reproducible PCR amplification of long stretches of DNA.

ACKNOWLEDGMENTS

We thank Henry Erlich, Tom Myers, and Rick Abramson (Cetus Corp.) for their helpful discussions. We thank Will Bloch (Cetus Corp.) for providing us with access to prototype AmpliWax PCR Gems developed by Will and members of his laboratory. We are grateful to Stan Rose and Bob Board (Perkin-Elmer Ce-

tus) for their support and excellent critique of this work and to David Gelfand (Cetus Corp.) for access to *Tma* and AmpliTaq Exo⁻ Mutein DNA polymerases. Doug Berg (Washington University) is thanked for providing the phage isolates carrying Tn5^{supF} insertions. Rajendra Krishnan (Washington University) is gratefully acknowledged for helping to get us going in the transposon business. We thank Judit Csejty (Cetus Corp.) for synthesis of all our primers, J. Fenton Williams (Perkin-Elmer Cetus) for his help in determining optimal thermocycling conditions using the GeneAmp PCR System 9600, Janet Ziegler (Applied Biosystems, Inc.) for assistance with, and suggestions for, adapting the capabilities of the ABI Gene Scanner to our particular uses, Pete Ohler for continuous computer assistance and graphics, and Eric Ladner for graphics and artwork.

REFERENCES

1. Jeffreys, A.L., V. Wilson, R. Neumann, and J. Keyte. 1988. Amplification of human mini-satellites by the polymerase chain reaction: Towards DNA fingerprinting of single cells. *Nucleic Acids Res.* **16**: 10953–10971.
2. Rychlik, W., W.J. Spencer, and R.E. Rhoads. 1990. Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucleic Acids Res.* **18**: 6409–6412.
3. Maga, E.A. and T. Richardson. 1991. Amplification of a 9.0 kb fragment using PCR. *BioTechniques* **11**: 185–186.
4. Krishnan, B.R., D. Kersulyte, I. Brikun, C.M. Berg, and D.E. Berg. 1991. *Nucleic Acids Res.* **19**: 6177–6182.
5. Erlich, H., D. Gelfand, and J. Sninsky. 1991. Recent advances in the polymerase chain reaction. *Science* **252**: 1643–1650.
6. Phadnis, S.H., H.V. Huang, and D.E. Berg. 1989. Tn5^{supF}, a 264 base pair transposon derived from Tn5 for insertion mutagenesis and sequencing DNAs cloned in phage lambda. *Proc. Natl. Acad. Sci.* **86**: 5908–5912.
7. Chou, Q., M. Russell, D. Birch, J. Raymond, and W. Bloch, in preparation.
8. Rose, E.A. 1991. Applications of the polymerase chain reaction to genome analysis. *FASEB J.* **5**: 46–54.
9. Olson, M., L. Hood, C. Cantor, and D. Botstein. 1989. A common language for physical mapping of the human genome. *Science* **245**: 1434–1435.
10. Nelson, D.L., S.A. Ledbetter, L. Corbo, M.F. Victoria, R. Ramirez-Solis, T. Webster, D.H. Ledbetter, and C.T. Caskey. 1989. Alu polymerase chain reaction: A method for rapid isolation of human specific sequences from complex DNA sources. *Proc. Natl. Acad. Sci.* **86**: 6686–6690.
11. Green, E.D. and M.V. Olson. 1990. Systematic screening of yeast artificial chromosome libraries by use of the polymerase chain reaction. *Proc. Natl. Acad. Sci.* **87**: 1213–1217.
12. Riley, J., D. Ogilvie, R. Finnear, D. Jenner, S. Powell, R. Anand, J.C. Smith, and A.F. Markham. 1990. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res.* **18**: 2887–2890.
13. Corbo, L., J.A. Maley, D.L. Nelson, and C.T. Caskey. 1990. Direct cloning of human transcribed sequences using hn-RNA from hybrid cell lines. *Science* **249**: 652–659.
14. Gyllensten, U.B. and H.A. Erlich. 1988. Generation of single stranded DNA by the polymerase chain reaction and its application to direct sequencing of the HLA DQA locus. *Proc. Natl. Acad. Sci.* **85**: 7652–7656.
15. Innis, M., K.B. Myambo, D.H. Gelfand, and M.A.D. Brow. 1988. DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction amplified DNA. *Proc. Natl. Acad. Sci.* **85**: 9436–9440.
16. Maniatis, T., E.F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: A laboratory manual*, p. 443. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
17. Mullis, K.B. and F. Faloona. 1987. Specific synthesis of DNA *in vitro* via a polymerase catalyzed chain reaction. *Methods Enzymol.* **1155**: 335–350.
18. Mullis, K.B. 1991. The polymerase chain reaction in an anemic mode: How to avoid cold deoxyribonuclear fusion. *PCR Methods Applic.* **1**: 1–4.
19. Innis, M.A. and D.H. Gelfand. 1990. Optimization of PCRs. In *PCR protocols: A guide to methods and applications* (ed. M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White), pp. 3–12. Academic Press, San Diego, CA.
20. rTth DNA polymerase package insert. 1991. Perkin-Elmer Cetus.
21. Rose, E.A. 1990. PCR and the Human Genome Project: The 1990 Cold Spring Harbor Meeting on Genome Mapping and Sequencing. *Amplifications* **5**: 1–10.
22. Ohler, L.D. and E.A. Rose. 1991. Use of long range PCR in genome mapping. *Cytogenet. Cell Genet.* (in press).
23. Ohler, L.D. and E.A. Rose. 1991. Transposon-based mapping and sequencing of chromosome 11p13. *Cytogenet. Cell Genet.* (in press).
24. Strausbaugh, L.D., M.T. Bourke, M.T. Sommer, M.E. Coon, and C.M. Berg. 1990. Probe mapping to facilitate transposon based DNA sequencing. *Proc. Natl. Acad. Sci.* **86**: 5908–5912.
25. Myers, T.W. and D.H. Gelfand. 1991. Reverse transcription and DNA amplification by a *Thermus thermophilus* DNA polymerase. *Biochemistry* **30**: 7661–7666.
26. Abramson, R., Cetus Corp., personal communication.
27. Gelfand, D.H., Cetus Corp., personal communication.
28. AmpliTaq DNA Polymerase package insert. 1991. Perkin-Elmer Cetus.
29. Hot Tub DNA Polymerase package insert. 1991. Amersham.
30. VENT DNA Polymerase package insert. 1991. New England Biolabs.

Received December 30, 1991; accepted in revised form March 20, 1992.