



On beyond classic RACE (rapid amplification of cDNA ends).

M A Frohman

Genome Res. 1994 4: S40-S58

References This article cites 36 articles, 5 of which can be accessed free at:
<http://genome.cshlp.org/content/4/1/S40.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](#).

A promotional banner for Cellecta. On the left, the text reads "CRISPR and RNAi Genetic Screening. Your new superpower." In the center is a white button with the text "LEARN MORE". On the right is a woman in a red and white superhero costume with a red mask, and the Cellecta logo (a green molecular structure) with the word "CELLECTA" below it.

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

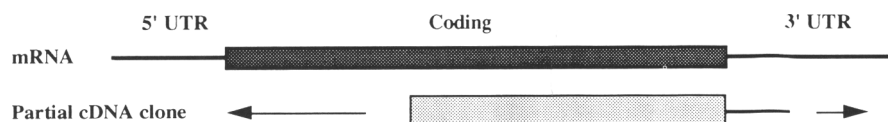
Copyright © Cold Spring Harbor Laboratory Press

On Beyond Classic RACE (Rapid Amplification of cDNA Ends)

Michael A. Frohman

Department of Pharmacology,
University Medical Center at Stony
Brook, Stony Brook,
New York 11794-8651

Most attempts to identify and isolate a novel cDNA result in the acquisition of clones that represent only a part of the mRNA's complete sequence. The missing sequence (cDNA ends) can be cloned by PCR, using a technique variously called rapid amplification of cDNA ends (RACE),⁽¹⁾ anchored PCR,⁽²⁾ or one-sided PCR.⁽³⁾ Since the initial reports of this technique, many laboratories have developed significant improvements on the basic approach.⁽⁴⁻¹⁶⁾ I will describe here the most recent hybrid version of the relatively simple classic RACE, and a more powerful but technically more challenging new RACE protocol that is adapted from the work of a number of laboratories.⁽¹⁷⁻²⁴⁾ Commercial RACE kits that are convenient but not as powerful as the most recent versions of classic and new RACE are available from Bethesda Research Laboratories⁽¹¹⁾ and Clontech.



Shown is a schematic representation of the setting in which RACE is useful in cDNA cloning strategies. Depicted is an mRNA for which a cDNA representing only an internal portion of the transcript has been obtained. Such circumstances often arise; one such example is when closely related genes are cloned using PCR amplification with degenerate primers encoding sequences homologous to amino acids found in all known members of the gene family.

OVERVIEW

Why use PCR (RACE) at all instead of screening (additional) cDNA libraries? RACE cloning is advantageous for several reasons: First, it takes weeks to screen cDNA libraries, obtain individual cDNA clones, and analyze the clones to determine whether the missing sequence is present. Using PCR, such information can be generated within a few days. As a result, it becomes practical to modify RNA preparation and/or reverse transcription conditions until full-length cDNAs are generated and observed. In addition, essentially unlimited numbers of independent clones can be generated using RACE, unlike library screens in which generally a single to a few cDNA clones are recovered. The availability of large numbers of clones provides confirmation of nucleotide sequence and allows the isolation of unusual transcripts that are alternately spliced or that begin at infrequently used promoters.

PRINCIPLES

Classic RACE

PCR is used to amplify partial cDNAs representing the region between a single point in a mRNA transcript and its 3' or 5' end (Fig. 1). A short internal stretch of sequence must already be known from the mRNA of interest. From this sequence, gene-specific primers are chosen that are oriented in the direction of the missing sequence. Extension of the partial cDNAs from the unknown end of the message back to the known region is achieved using primers that anneal to the preexisting poly(A) tail (3' end) or an appended homopolymer tail (5' end). Using RACE, enrichments on the order of 10^6 to 10^7 -fold can be obtained. As a result, relatively pure cDNA "ends" are generated that can be easily cloned or rapidly characterized using conventional techniques.⁽¹⁾

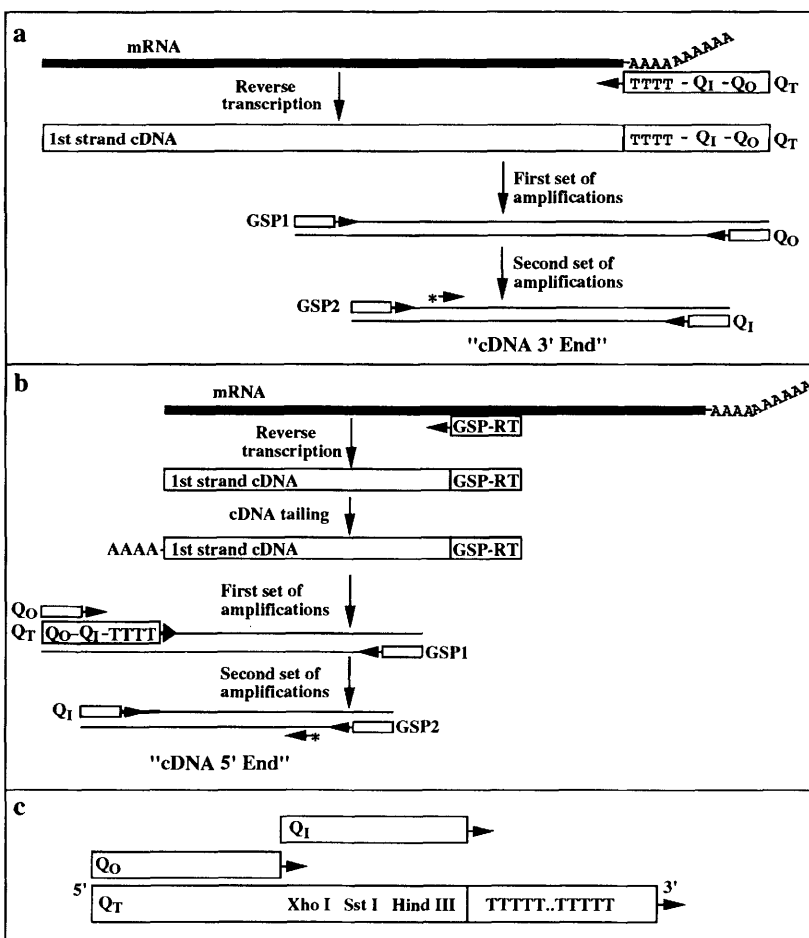


FIGURE 1 Schematic representation of classic RACE. Explanations are given in the text. At each step, the diagram is simplified to illustrate only how the new product formed during the previous step is utilized. (GSP1) Gene-specific primer 1; (GSP2) gene-specific primer 2; (GSP-RT) gene-specific primer used for reverse transcription; (*→)GSP-Hyb/Seq or gene-specific primer for use in hybridization and sequencing reactions. (a) Amplification of 3' partial cDNA ends; (b) amplification of 5' partial cDNA ends; (c) schematic representation of the primers used in classic RACE. The 52-nucleotide Q_T primer (5' Q_O - Q_I -TTTT 3') contains a 17-nucleotide oligo(dT) sequence at the 3' end followed by a 35-nucleotide sequence encoding *Hind*III, *Sst*I, and *Xho*I recognition sites. The Q_I and Q_O primers overlap by 1 nucleotide; the Q_I primer contains all three of the recognition sites.

Primers:

Q_T : 5'-CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGCTTTTTTTTTTTTTTTTTT-3'

Q_O : 5'-CCAGTGAGCAGAGTGACG-3'

Q_I : 5'-GAGGACTCGAGCTCAAGC-3'

To generate "3' end" partial cDNA clones, mRNA is reverse transcribed using a "hybrid" primer (Q_T) that consists of 17 nucleotides of oligo(dT) followed by a unique 35-base oligonucleotide sequence (Q_I - Q_O ; Fig. 1a and c), which in many reports is denoted as an "anchor" primer. Amplification is then performed using a primer containing part of this sequence (Q_O) that now binds to each cDNA at its 3' end, and using a primer derived from the gene of interest (GSP1). A second set of amplification cycles is then carried out using "nested" primers (Q_I and GSP2) to quench the amplification of nonspecific products. To generate "5'-end" partial cDNA clones, reverse transcription (primer extension) is carried out using a gene-specific primer (GSP-RT; Fig. 1b) to generate first-strand products. Then, a poly(A) tail is appended

using terminal deoxynucleotidyltransferase (TdT) and dATP. Amplification is then achieved using (1) the hybrid primer Q_T to form the second strand of cDNA, (2) the Q_O primer, and (3) a gene-specific primer upstream of the one used for reverse transcription. Finally, a second set of PCR cycles is carried out using nested primers (Q_I and GSP2) to increase specificity.⁽⁵⁾

Classic RACE Variations

In general, as described above, the gene-specific primer is derived from a short stretch of sequence that is already known from the mRNA of interest. A frequent question is whether degenerate primers, that is, those directed against predicted nucleotide sequence based on known amino acid sequence, can be used instead. Although such primers increase the quantity of spurious amplification, the approach can work, if other parameters are favorable (i.e., message abundance, GC composition, and cDNA end size; see Monstein et al.).⁽¹⁴⁾

At the unknown end of the cDNA, the 5' end can be tailed with C's instead of A's, and then amplified using a hybrid primer with a tail containing G's,⁽²⁾ or a mixture of G's and inosines (I).⁽¹¹⁾ Although the G:I approach entails synthesizing a primer that can be used for 5' RACE only [because a T-tailed primer must be used to anneal to the poly(A) tail of the 3' end], there may be sufficient benefits from using a mixed G:I tail to justify the cost, as the G:I region should anneal at temperatures similar to those of other primers normally used in PCR. In contrast, it is believed that homopolymers of either T's or G's present problems during PCR, because of the very low and very high annealing temperatures, respectively, required for their optimal usage.^(1,11) On the other hand, the inosine residues function as degenerate nucleotides and will lead to higher spurious amplification, so the magnitude of the benefit of using a mixed G:I primer is unknown.

To minimize the length of homopolymer tail actually amplified, a "Lock-Docking" primer was developed by Borson et al.⁽⁸⁾ In this approach, the final 2 nucleotides on the 3' end of the primer are degenerate. For example, to amplify cDNAs linked to an A-tail, the lock-docking primer would look like



where X represents (e.g.) one or more restriction sites at the 5' end of the primer. The advantage of this approach is that it forces the primer to anneal to the junction of the natural or appended homopolymer tail and the cDNA sequence. The disadvantage is that it is necessary to synthesize four primers, as most synthesizers can only synthesize primers starting from an unambiguous 3' end.

In another variation, the location of the anchor primer is changed from the end of the unknown region of sequence to random points within the unknown region.⁽⁷⁾ This is accomplished using a primer containing an anchor region followed by six random nucleotides (5'-XXXXXX-NNNNNN-3') either for reverse transcription (3' RACE) or for creation of the second strand of cDNA (5' RACE). This approach is valuable when the 3' or 5' ends lie so far away from the region of known sequence that the entire unknown region cannot be amplified effectively. Using this approach, cDNA ends of defined sizes are not generated; instead, a library of randomly sized fragments is

obtained, all of which initiate at the gene-specific primer. The largest fragments can be cloned and characterized, extending the length of the known sequence, and the process (or standard RACE) repeated until the real unknown end is identified. The development of "long" PCR may make this approach unnecessary.

New RACE

The most technically challenging step in classic 5' RACE is to cajole reverse transcriptase to copy the mRNA of interest in its entirety into first-strand cDNA. Because prematurely terminated first-strand cDNAs are tailed by terminal transferase just as effectively as full-length cDNAs, cDNA populations composed largely of prematurely terminated first strands will result primarily in the amplification and recovery of cDNA ends that are not full length either (Fig. 2a). This problem is encountered routinely for vertebrate genes, which are often quite GC rich at their 5' ends and thus frequently contain sequences that hinder reverse transcription. A number of laboratories have developed steps or protocols designed to approach the problem,⁽¹⁷⁻²⁴⁾ the protocol described here and denoted "new RACE" is for the most part a composite adapted from the cited reports.

New RACE departs from classic RACE in that the "anchor" primer is attached to the 5' end of the mRNA **before** the reverse transcription step. Hence, the anchor sequence becomes incorporated into the first-strand cDNA if the reverse transcription proceeds through the entire length of the mRNA of interest (and through the relatively short anchor sequence) as shown in Figure 2b.

Before beginning new RACE (Fig. 3a), the mRNA is subjected to a dephos-

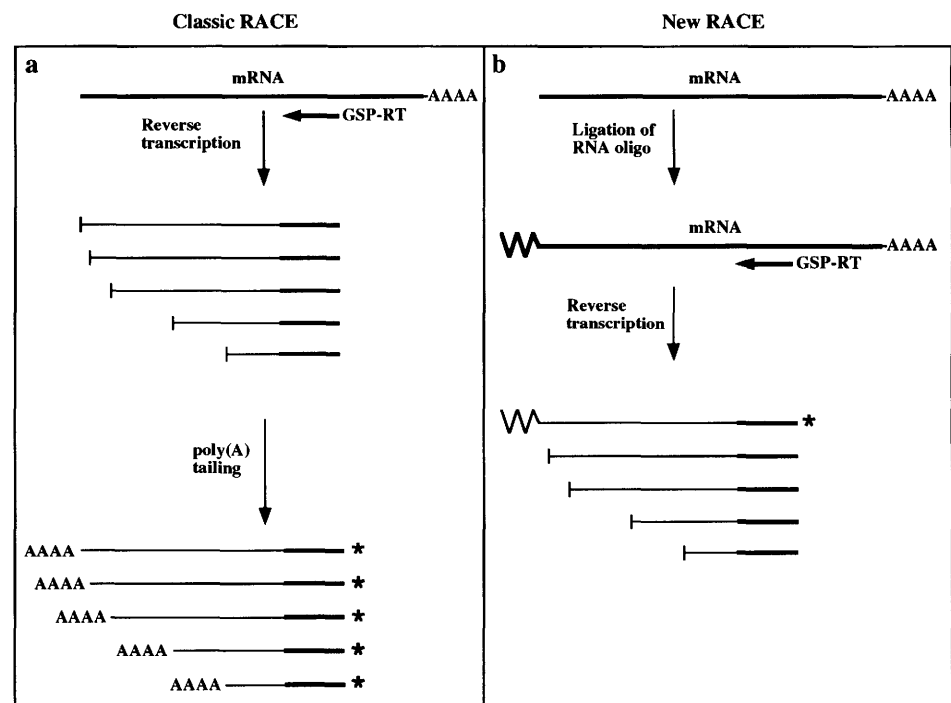


FIGURE 2 A depiction of the advantage of using new RACE over classic RACE. (a) In classic RACE, premature termination in the reverse transcription step results in polyadenylation of less than full-length first-strand cDNAs, all of which can be amplified using PCR to generate less than full-length cDNA 5' ends. (*) cDNA ends created that will be amplified in the subsequent PCR reaction. (b) In new RACE, less than full-length cDNAs are also created but are not terminated by the anchor sequence; hence, they cannot be amplified in the subsequent PCR reaction.

phorylation step using calf intestinal phosphatase (CIP). This step actually does nothing to full-length mRNAs, which have methyl-G caps at their termini; but it does dephosphorylate degraded mRNAs, which are uncapped at their termini.⁽¹⁹⁾ This makes the degraded RNA biologically inert during the ensuing ligation step, because the phosphate group is required to drive the reaction. The full-length mRNAs are then decapped using tobacco acid pyrophosphatase (TAP), which leaves them with an active and phosphorylated 5' terminus.^(18,22) Using T4 RNA ligase, this mRNA is then ligated to a short synthetic RNA oligonucleotide that has been generated by in vitro transcription of a linearized plasmid (Fig. 3b).⁽¹⁷⁾ The RNA oligonucleotide mRNA hybrids are then reverse transcribed using a gene-specific primer or random primers to create first-strand cDNA. Finally, the 5' cDNA ends are amplified in two nested PCR reactions using additional gene-specific primers and primers derived from the sequence of the RNA oligonucleotide.

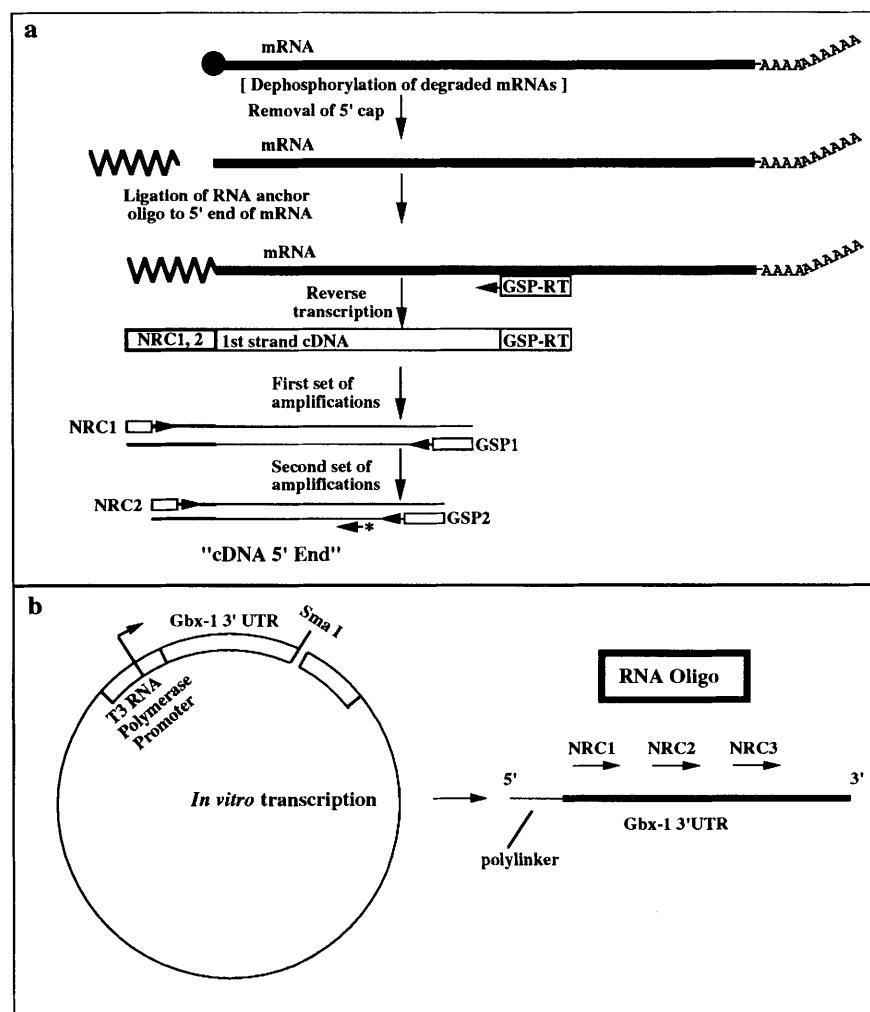


FIGURE 3 Schematic representation of new RACE. Explanations are given in the text. At each step, the diagram is simplified to illustrate only how the new product formed during the previous step is utilized. See the legend to Fig. 1 for description of some primers. (a) Amplification of 5' partial cDNA ends; (b) In vitro synthesis of the RNA oligonucleotide used for ligation in new RACE and schematic representation of the corresponding required primers. A 132-nucleotide RNA oligonucleotide is produced by in vitro transcription of the plasmid depicted using T3 RNA polymerase. Primers NRC-1, NRC-2, and NRC-3 are derived from the sequence of the oligonucleotide but do not encode restriction sites. To assist in the cloning of cDNA ends, the sequence "ATCG" is added to the 5' end of NRC-2, as described in the cloning section of the text.

The new RACE approach can also be used to generate 3' cDNA ends⁽¹⁹⁾ (see related protocols),^(18,20) and is useful in particular for nonpolyadenylated RNAs. In brief, cytoplasmic RNA is dephosphorylated and ligated to a short synthetic RNA oligonucleotide as described above. Although ligation of the oligonucleotide to the 5' end of the RNA was emphasized above, RNA oligonucleotides actually ligate to both ends of the cytoplasmic RNAs. For the reverse transcription step, a primer derived from the RNA oligonucleotide sequence is used (e.g., the reverse complement of NRC-3, Fig. 3). Reverse transcription of the RNA oligonucleotides that happen to be ligated to the 3' end of the cytoplasmic RNAs results in the creation of cDNAs that have the RNA oligonucleotide sequence appended to their 3' end. Gene-specific primers oriented in the 5' → 3' direction and new RACE primers (e.g., the reverse complements of NRC-2 and NRC-1; Fig. 3) can be used in nested PCR reactions to amplify the 3' ends.

Protocols

CLASSIC RACE

Materials

The materials required for this procedure can be purchased, along with the appropriate 5× or 10× enzyme reaction buffers, from most major suppliers. MMLV reverse transcriptase can be obtained from Bethesda Research Labs (BRL), RNasin and *Taq* polymerase from Promega Biotech, and TdT from either BRL or Boehringer Mannheim. Enzymes are used as directed by the suppliers, except for *Taq* polymerase. Instead of using the recommended reaction mixture, a 10× buffer consisting of 670 mM Tris-HCl (pH 9.0), 67 mM MgCl₂, 1700 μg/ml of BSA, and 166 mM (NH₄)₂SO₄ is substituted, and reaction conditions are altered as described further below.^(1,16) Oligonucleotide primer sequences are listed in the legend to Figure 1. Primers can be used "crude" except for Q_T, which should be purified to ensure that it is uniformly full length. dNTPs can be purchased as 100 mM solutions from PL-Biochemicals/Pharmacia or Boehringer Mannheim.

3' End cDNA amplification

Step 1. Reverse Transcription to Generate cDNA Templates Procedure: Assemble reverse transcription components on ice: 4 μl of 5× reverse transcription buffer, 1.3 μl of dNTPs (stock concentration is 15 mM of each dNTP), 0.25 μl (10 units) of RNasin, and 0.5 μl of Q_T primer (100 ng/μl). Heat 1 μg of poly(A)⁺/RNA or 5 μg of total RNA in 14 μl of water at 80°C for 3 min and cool rapidly on ice. Add to reverse transcription components. Add 1 μl (200 units) of MMLV reverse transcriptase, and incubate for 1 hr at 37°C, 30 min at 42°C, and 10 min at 50°C. Dilute the reaction mixture to 1 ml with TE [10 mM Tris-HCl (pH 7.5)/1 mM EDTA] and store at 4°C (3' end cDNA pool).

Comments: Poly(A)⁺ RNA is used preferentially for reverse transcription to decrease background, but it is unnecessary to prepare it if only total RNA is available. An important factor in the generation of full-length 3' end partial cDNAs concerns the stringency of the reverse transcription reaction. Reverse transcription reactions were historically carried out at relatively low temperatures (37–42°C) using a vast excess of primer (~½ the mass of the mRNA, which represents an ~30:1 molar ratio). Under these low stringency conditions, a stretch of A residues as short as 6–8 nucleotides will suffice as a binding site for an oligo(dT)-tailed primer. This may result in cDNA synthesis

being initiated at sites upstream of the poly(A) tail, leading to truncation of the desired amplification product. It should be assumed that this has occurred if a canonical polyadenylation signal sequence⁽²⁵⁾ is not found near the 3' end of the cDNAs generated. This can be minimized by controlling two parameters: primer concentration and reaction temperature. The primer concentration can be reduced dramatically without decreasing significantly the amount of cDNA synthesized⁽²⁶⁾ and will begin to bind preferentially to the longest A-rich stretches present [i.e., the poly(A) tail]. The quantity recommended above represents a good starting point; it can be reduced fivefold further if significant truncation is observed.

In the protocol described above, the incubation temperature is raised slowly to encourage reverse transcription to proceed through regions of difficult secondary structure. Because the half-life of MMLV rapidly decreases as the incubation temperature increases, the reaction cannot be carried out at elevated temperatures in its entirety. Alternatively, the problem of difficult secondary structure (and nonspecific reverse transcription) can be approached using heat stable reverse transcriptases, which are now available from several suppliers (Perkin-Elmer Cetus, Amersham, Epicentre Technologies, and others). As in PCR reactions, the stringency of reverse transcription can thus be controlled by adjusting the temperature at which the primer is annealed to the mRNA. The optimal temperature depends on the specific reaction buffer and reverse transcriptase used and should be determined empirically but will usually be found to be in the range of 48–56°C for a primer terminated by a 17-nucleotide oligo(dT) tail.

Step 2. Amplification Procedure: First round: Add an aliquot of the cDNA pool (1 μ l) and primers (25 pmoles each of GSP1 and Q_O) to 50 μ l of PCR cocktail (1 \times *Taq* polymerase buffer, as described above, each dNTP at 1.5 mM and 10% DMSO) in a 0.5-ml microcentrifuge tube, and heat in a DNA thermal cycler for 5 min at 98°C to denature the first-strand products. Cool to 75°C. Add 2.5 units of *Taq* polymerase, overlay the mixture with 30 μ l of mineral oil (Sigma 400-5; preheat the oil in the thermal cycler to 75°C, and incubate at the appropriate annealing temperature (52–60°C) for 2 min. Extend the cDNAs at 72°C for 40 min. Carry out 30 cycles of amplification using a step program (94°C, 1 min; 52–60°C, 1 min; 72°C, 3 min), followed by a 15-min final extension at 72°C. Cool to room temperature.

Second round: Dilute a portion of the amplification products from the first round 1:20 in TE. Amplify 1 μ l of the diluted material with primers GSP2 and Q_I using the procedure described above, but eliminate the initial 2-min annealing step and the 72°C, 40-min extension step.

Comments: It is important to add the *Taq* polymerase *after* heating the mixture to a temperature above the T_m of the primers (hot start PCR). Addition of the enzyme prior to this point allows one cycle to take place at room temperature, promoting the synthesis of nonspecific background products dependent on low-stringency interactions.

An annealing temperature close to the effective T_m of the primers should be used. The Q_I and Q_O primers work well at 60°C under the PCR conditions recommended here, although the actual optimal temperature may depend on the PCR machine used. Gene-specific primers of similar length and GC content should be chosen. Computer programs to assist in the selection of primers are widely available and should be used. An extension time of 1 min/kb expected product should be allowed during the amplification cycles. If the expected length of product is unknown, try 3–4 min initially.

Very little substrate is required for the PCR reaction: 1 μ g of poly(A)⁺ RNA typically contains $\sim 5 \times 10^7$ copies of each low abundance transcript. The PCR reaction described here works optimally when 10^3 – 10^5 templates (of the de-

sired cDNA) are present in the starting mixture; thus, as little as 0.002% of the reverse transcription mixture suffices for the PCR reaction. Addition of too much starting material to the amplification reaction will lead to production of large amounts of nonspecific product and should be avoided. The RACE technique is particularly sensitive to this problem, as every DNA in the mixture, desired and undesired, contains a binding site for the Q_0 and Q_1 primers.

It was found empirically that allowing extra extension time (40 min) during the first amplification round (when the second strand of cDNA is created) sometimes resulted in increased yields of the specific product relative to background amplification, and in particular, increased the yields of long cDNAs versus short cDNAs when specific cDNA ends of multiple lengths were present.⁽¹⁾ Prior treatment of cDNA templates with RNA hydrolysis or a combination of RNase H and RNase A infrequently improves the efficiency of amplification of specific cDNAs.

Choice of buffer systems: For some applications intended for cloned PCR products, such as expressing cDNAs to generate proteins, it is critically important to minimize the rate at which mutations occur during amplification. In other applications, such as using the cloned DNA as a probe in hybridization experiments, the presence of a few mutations is relatively unimportant and thus it is most convenient to use PCR conditions that maximize the likelihood of generating the desired product the first time a set of primers is used. Unfortunately, PCR conditions that result in a minimum of mutations can result in erratic yields and often the desired product cannot be generated until the PCR conditions have been optimized, whereas PCR conditions that reliably produce desired products result in a relatively high mutation rate (~1% after 30 rounds). Thus, appropriate conditions must be chosen to generate the PCR products required prior to undertaking cloning steps. PCR conditions that result in a minimum of mutations require the use of nucleotides (dNTPs) at low concentrations (0.2 mM). Using the conditions recommended for *Taq* polymerase by Perkin-Elmer Cetus results in an error rate of ~0.05% after 30 rounds of amplification. However, the conditions recommended often have to be optimized, meaning that the pH of the buffer and the concentration of magnesium have to be adjusted until the desired product is observed. In addition, inclusion of DMSO or formamide may be required. For those who do not wish to prepare their own reagents to carry out optimization experiments, such kits are commercially available (e.g., from Invitrogen and Stratagene). PCR conditions that work much more frequently in the absence of optimization steps require the use of DMSO, ammonium sulfate, and relatively high concentrations (1.5 mM) of dNTPs, as described above. It should be noted that the inclusion of DMSO to 10% decreases primer melting temperatures (and thus optimal annealing temperatures) by ~5–6°C.

5' End cDNA Amplification

Step 1. Reverse Transcription to Generate cDNA Templates Procedure: Assemble reverse transcription components on ice: 4 μ l of 5 \times reverse transcription buffer, 1.3 μ l of dNTPs (stock concentration is 15 mM of each dNTP), 0.25 μ l (10 units) of RNasin, and 0.5 μ l of GSP-RT primer (100 ng/ μ l). Heat 1 μ g of poly(A)⁺ RNA or 5 μ g of total RNA in 14 μ l of water at 80°C for 3 min and cool rapidly on ice. Add to reverse transcription components. Add 1 μ l (200 units) of MMLV reverse transcriptase and incubate for 1 hr at 37°C, 30 min at 42°C, and 10 min at 50°C. Dilute the reaction mixture to 400 μ l with TE [10 mM Tris-HCl (pH 7.5)/1 mM EDTA] and store at 4°C (5'-end nontailed cDNA pool).

Comments: Many of the remarks made above in the section on reverse transcribing 3' end partial cDNAs are also relevant here and should be noted.

There is, however, one major difference. The efficiency of cDNA extension is now critically important, as each specific cDNA, no matter how short, is subsequently tailed and becomes a suitable template for amplification (Fig. 2a). Thus, the PCR products eventually generated directly reflect the quality of the reverse transcription reaction. Extension can be maximized by using clean, intact RNA, by selecting the primer for reverse transcription to be near the 5' end of region of known sequence, and in theory by using heat-stable reverse transcriptase at elevated temperatures or a combination of MMLV and heat-stable reverse transcriptase at multiple temperatures. Synthesis of cDNAs at elevated temperatures should diminish the amount of secondary structure encountered in GC-rich regions of the mRNA.

Step 2. Appending a Poly(A) Tail to First-Strand cDNA Products Procedure: Remove excess primer using Microcon-100 spin filters (Amicon Corp.) or an equivalent product, following the manufacturer's instructions. Wash the material by spin filtration twice more using TE. The final volume recovered should not exceed 10 μ l.

Add 4 μ l of 5 \times tailing buffer [125 mM Tris-HCl (pH 6.6), 1 M KAcodylate, and 1250 μ g/ml BSA], 1.2 μ l of 25 mM CoCl₂, 4 μ l of 1 mM dATP, and 10 units of TdT. Incubate for 5 min at 37°C and then 5 min at 65°C. Dilute to 500 μ l with TE (5'-end-tailed cDNA pool).

Comments: To attach a known sequence to the 5' end of the first strand cDNA, a homopolymeric tail is appended using TdT. We prefer appending poly(A) tails rather than poly(C) tails for several reasons. First, the 3'-end strategy is based on the naturally occurring poly(A) tail; thus, the same adapter primers can be used for both ends, decreasing variability in the protocol and cost. Second, because A/T binding is weaker than G/C binding, longer stretches of A residues ($\sim 2\times$) are required before the oligo(dT)-tailed Q_T primer will bind to an internal site and truncate the amplification product. Third, vertebrate coding sequences and 5' untranslated regions tend to be biased toward G/C residues; thus, use of a poly(A) tail further decreases the likelihood of inappropriate truncation.

Unlike many other situations in which homopolymeric tails are appended, the actual length of the tail added here is unimportant, as long as it exceeds 17 nucleotides. This is because although the oligo(dT)-tailed primers subsequently bind all along the length of the appended poly(A) tail, only the innermost one becomes incorporated into the amplification product, and consequently, the remainder of the poly(A) tail is lost.⁽¹⁾ The truncation appears to happen because *Taq* polymerase is unable to resolve branched structures (efficiently). The conditions described in the procedure above result in the addition of 30–400 nucleotides.

Step 3. Amplification Procedure: First round: Add an aliquot of the 5' end tailed cDNA pool (1 μ l) and primers (25 pmoles each of GSP1 and Q_O, as shown in Fig. 2b; and 2 pmoles of Q_T) to 50 μ l of PCR cocktail (1 \times *Taq* polymerase buffer, each dNTP at 1.5 mM, and 10% DMSO) in a 0.5-ml microcentrifuge tube, and heat in a DNA thermal cycler for 5 min at 98°C to denature the first-strand products. Cool to 75°C. Add 2.5 units of *Taq* polymerase, overlay the mixture with 30 μ l of mineral oil (Sigma 400-5; preheat it in the thermal cycler to 75°C), and incubate at the appropriate annealing temperature (48–52°C) for 2 min. Extend the cDNAs at 72°C for 40 min. Carry out 30 cycles of amplification using a step program (94°C, 1 min; 52–60°C, 1 min; 72°C, 3 min), followed by a 15-min final extension at 72°C. Cool to room temperature.

Second round: Dilute a portion of the amplification products from the first-round 1:20 in TE. Amplify 1 μ l of the diluted material with primers GSP2

and Q₁ using the procedure described above, but eliminate the initial 2-min annealing step and the 72°C, 40-min extension step.

Comments: Many of the remarks made above in the section on amplifying 3' end partial cDNAs are also relevant here and should be noted. There is, however, one major difference. The annealing temperature in the first step (48–52°C) is lower than that used in successive cycles (52–60°C). This is because cDNA synthesis during the first round depends on the interaction of the appended poly(A) tail and the oligo(dT)-tailed Q_T primer, whereas in all subsequent rounds, amplification can proceed using the Q_O primer, which is composed of ~60% GC and can anneal at a much higher temperature to its complementary target.

NEW RACE

Materials

See the Classic RACE (above) for sources of some of the required materials. TAP and a cost-effective RNA transcription kit can be purchased along with the appropriate 10× enzyme reaction buffers from Epicentre (Madison, WI). T4 RNA ligase can be purchased from New England Biolabs or Boehringer Mannheim. Note that 10× T4 RNA ligase buffers supplied by some manufacturers contain too much ATP.⁽¹⁷⁾ Check the composition of any commercially supplied 10× buffer and make your own if it contains more than 1 mM ATP (final 1× concentration should be 0.1 mM), as described below.

5' End cDNA amplification

The procedure detailed below is described using relatively large amounts of RNA and can be scaled down if RNA quantities are limited. The advantage of starting with large amounts of RNA is that aliquots can be electrophoresed quickly after each step of the procedure to confirm that detectable degradation of the RNA has not occurred, and the dephosphorylated, decapped, ligated RNA can be stored indefinitely for many future experiments.

Step 1. Dephosphorylation of degraded RNAs Procedure: In general, follow the manufacturer's recommendations for use of the phosphatase. To dephosphorylate RNA using CIP from Boehringer Mannheim, prepare a reaction mixture containing (50 μg of RNA in 41 μl of water, 5 μl of 10× buffer, 0.5 μl of 100 mM DTT, 1.25 μl of RNasin (40 U/μl), and 3.5 μl of CIP (1 U/μl). Incubate the reaction at 50°C for 1 hr, and add proteinase K (Boehringer Mannheim) to 50 μg/ml and incubate at 37°C for 30 min. Extract the reaction with a mixture of phenol/chloroform, extract again with chloroform, and precipitate the RNA using 1/10 volume of 3 M NaOAc and 2.5 volumes of ethanol. Resuspend the RNA in 43.6 μl of water. Electrophorese 2 μg (1.6 μl) on a 1% TAE agarose gel adjacent to a lane containing 2 μg of the original RNA preparation, stain the gel with ethidium bromide, and confirm visually that the RNA remained intact during the dephosphorylation step.

Step 2. Decapping of intact RNAs Procedure: Prepare a reaction mixture containing 38 μg of RNA in 42 μl of water (this is the RNA recovered from step 1), 5 μl of 10× TAP buffer; 1.25 μl of RNasin, (40 U/2μl), 1 μl of 100 mM ATP; and 1 μl of TAP (5 U/μl). Incubate the reaction at 37°C for 1 hr and add 200 μl of TE. Extract the reaction with a mixture of phenol/chloroform, extract again with chloroform, and precipitate the RNA using 1/10 volume of 3 M NaOAc and 2.5 volumes of ethanol. Resuspend the RNA in 40 μl of water.

Electrophorese 2 μg on a TAE 1% agarose gel adjacent to a lane containing 2 μg of the original RNA preparation, stain the gel with ethidium bromide, and confirm visually that the RNA remained intact during the decapping step. **Comment:** Most protocols call for much more TAP. The excess is not necessary, and the enzyme is very expensive.

Step 3. Preparation of RNA oligonucleotides Choose a plasmid that can be linearized at a site ~ 100 bp downstream from a T7 or T3 RNA polymerase site (see Fig. 3b). Ideally, a plasmid containing some insert cloned into the first polylinker site is optimal, because primers made from palindromic polylinker DNA do not perform well in PCR. For my experiments, I use the 3' UTR of the mouse gene Gbx-1,⁽²⁷⁾ which is cloned into the *Sst*I site of pBS-SK (Stratagene). I linearize with *Sma*I and transcribe with T3 RNA polymerase to produce a 132-nucleotide RNA oligonucleotide, of which all but 17 nucleotides are from Gbx-1. Note that adenosines are the best acceptors for the 3' end of the RNA oligonucleotide to ligate to the 5' end of its target, if an appropriate restriction site can be found. The primers subsequently used for amplification are all derived from the Gbx-1 3' UTR sequence. Interested investigators are welcome to the Gbx-1 NRC primer sequences and plasmid upon request.

Carry out a test transcription to make sure that everything is working; then scale up. The oligonucleotide can be stored at -80°C indefinitely for many future experiments, and it is important to synthesize enough oligonucleotide so that losses due to purification and spot checks along the way will leave plenty of material at the end of the procedure.

Procedure: Linearize 25 μg of the plasmid that is to be transcribed (the plasmid should be reasonably free of RNases). Treat the digestion reaction with 50 $\mu\text{g}/\text{ml}$ of Proteinase K for 30 min at 37°C , followed by $2\times$ phenol/ CHCl_3 extractions, one CHCl_3 extraction, and ethanol precipitation. Resuspend the template DNA in 25 μl of TE (pH 8.0) for a final concentration of ~ 1 $\mu\text{g}/\mu\text{l}$.

For transcription, mix at room temperature in the following order, and then incubate the reaction mixture at 37°C for 1 hr.

	Test Scale	or	Prep Scale
DEPC-water	4 μl		80
5 \times buffer	2		40
0.1 M DTT	1		20
10 mM UTP	0.5		10
10 mM ATP	0.5		10
10 mM CTP	0.5		10
10 mM GTP	0.5		10
Restricted DNA (1 $\mu\text{g}/\mu\text{l}$)	0.5		10
RNasin (40 U/ μl)	0.25		5
RNA polymerase (20 U/ μl)	0.25		5

Add 0.5 μl DNase (RNase free) for every 20 μl of reaction volume and incubate at 37°C for 10 min. Run 5 μl of test or prep reaction on a 1% TAE agarose gel to check. Expect to see a diffuse band at about the right size (or a bit smaller) in addition to some smearing all up and down the gel.

Purify the oligo by extracting with phenol/ CHCl_3 and CHCl_3 , and then clean three times using water and a Microcon spin filter (prerinsed with water). Run another appropriately sized aliquot on a 1% TAE agarose gel to check integrity and concentration of oligo. Microcon 30 spin filters have a cut off size of 60 nucleotides, and Microcon 100 spin filters have a cut off size of 300 nucleotides. Microcon 10 spin filters are probably most appropriate if the oligo is <100 nucleotides, and Microcon 30 spin filters for anything larger.

Step 4. RNA Oligonucleotide–cellular RNA Ligation Set up two tubes: one with TAP_{ped} cellular RNA, the other with unTAP_{ped} cellular RNA.

11.25 μ l	water
3 μ l	10 \times buffer = 500 mM Tris (pH 7.9), 100 mM MgCl ₂ , 20 mM DTT, 1 mg/ml of BSA
0.75 μ l	RNasin (40 U/ μ l)
2 μ l	4 μ g of RNA oligonucleotide (3–6 molar excess over target cellular RNA)
10 μ l	10 μ g of TAP _{ped} (or unTAP _{ped}) RNA
1.5 μ l	2 mM ATP
1.5 μ l	T4 RNA ligase (20 U/ μ l)
30 μ l	

Incubate for 16 hr at 17°C. Purify the ligated oligonucleotide–RNA using Microcon 100 spin filtration (3 \times in water; prerinse filter with RNase-free water). The volume recovered should not exceed 20 μ l. Run one-third of the ligation on a 1% TAE agarose gel to check integrity of the ligated RNA. It should look about as it did before ligation.

Step 5. Reverse Transcription Procedure: Assemble reverse transcription components on ice: 4 μ l of 5 \times reverse transcription buffer, 1.3 μ l of dNTPs (stock concentration is 15 mM of each dNTP), 0.25 μ l (10 units) of RNasin, and 1 μ l of antisense-specific primer (20 ng/ μ l), or random hexamers (100 ng/ μ l). Heat the remaining RNA (~6.7 μ g) in 14 μ l of water at 80°C for 3 min, and cool rapidly on ice. Add to reverse transcription components. Add 1 μ l (200 units) of MMLV reverse transcriptase and incubate for 1 hr at 37°C, 30 min at 42°C, and 10 min at 50°C. Dilute the reaction mixture to 100 μ l with TE [10 mM Tris-HCl (pH 7.5)/1 mM EDTA] and store at 4°C (5' end oligonucleotide–cDNA pool).

Step 6. Amplification Procedure: First round: Add an aliquot of the 5' end oligonucleotide–cDNA pool (1 μ l) and primers (25 pmoles each of GSP1 and NRC-1, as in Fig. 3a) to 50 μ l of PCR cocktail (1 \times *Taq* polymerase buffer, each dNTP at 1.5 mM, and 10% DMSO) in a 0.5-ml microcentrifuge tube, and heat in a DNA thermal cycler for 5 min at 98°C to denature the first-strand products. Cool to 75°C. Add 2.5 units of *Taq* polymerase, overlay the mixture with 30 μ l of mineral oil (Sigma 400-5; preheat it in the thermal cycler to 75°C), and incubate at the appropriate annealing temperature (52–60°C) for 2 min. Extend the cDNAs at 72°C for 40 min. Carry out 35 cycles of amplification using a step program (94°C, 1 min; 52–60°C, 1 min; 72°C, 3 min), followed by a 15-min final extension at 72°C. Cool to room temperature.

Second round: Dilute a portion of the amplification products from the first round 1:20 in TE. Amplify 1 μ l of the diluted material with primers GSP2 and NRC-2 using the procedure described above, but eliminate the initial 2-min annealing step, and the 72°C, 40-min extension step.

Comments: Many of the remarks made in the Classic RACE (above) on amplifying 5'-end partial cDNAs are also relevant here and should be noted.

Analysis of Amplification Products

The production of specific partial cDNAs by the RACE protocol is assessed using Southern blot hybridization analysis. After the second set of amplification cycles, the first and second set reaction products are electrophoresed in a 1% agarose gel, stained with ethidium bromide (EtBr), denatured, and transferred to a nylon membrane. After hybridization with a labeled oligomer or gene fragment derived from a region contained within the amplified fragment (e.g., GSP-Hyb/Seq in Fig. 1a,b), gene-specific partial cDNA ends should be detected easily. Yields of the desired product relative to nonspecific am-

plified cDNA in the first round products should vary from <1% of the amplified material to nearly 100%, depending largely on the stringency of the amplification reaction, the amplification efficiency of the specific cDNA end, and the relative abundance of the specific transcript within the mRNA source. In the second set of amplification cycles, ~100% of the cDNA detected by EtBr staining should represent specific product. If specific hybridization is not observed, then troubleshooting steps should be initiated.

Information gained from this analysis should be used to optimize the RACE procedure. If low yields of specific product are observed because non-specific products are being amplified efficiently, then annealing temperatures can be raised gradually (~2°C at a time) and sequentially in each stage of the procedure until nonspecific products are no longer observed. Alternatively, some investigators have reported success using the touchdown PCR procedure to optimize the annealing temperature without trial and error.⁽²⁸⁾ Optimizing the annealing temperature is also indicated if multiple species of specific products are observed, which could indicate that truncation of specific products is occurring. If multiple species of specific products are observed after the reverse transcription and amplification reactions have been fully optimized, then the possibility should be entertained that alternate splicing or promoter use is occurring.

CLASSIC RACE ONLY

If a nearly continuous smear of specific products is observed up to a specific size limit after 5' end amplification, this suggests that polymerase pausing occurred during the reverse transcription step. To obtain nearly full-length cDNA ends, the amplification mixture should be electrophoresed and the longest products recovered by gel isolation. An aliquot of this material can then be reamplified for a limited number of cycles.

NEW RACE ONLY

Expect to see one or two extra nucleotides insert between the RNA oligonucleotide 3' end, and the 5' end of the gene of interest—these come from the transcription step using T7, T3, or SP6, which can add an extra nucleotide or two to oligonucleotides past the end of the template (template-independent transcription).

Compare the results from unTAP_{ped} RNA versus TAP_{ped} RNA. Junction sites (where the oligonucleotide is connected to the 5' end of your gene) in common arise from ligation of the oligonucleotide to degraded RNA; unique junctions in the TAP_{ped} RNA population represent candidate transcription start sites. If you have RNA degradation sites (e.g., TTT▼AAA) in your 5' RNA end, you may have substantial numbers of clones that begin at exactly the same nucleotide but arise from ligation of the oligonucleotide to degraded RNA molecules, not from ligation of the oligonucleotide to the true 5' end of the RNA.

Look for TATA, CCAAT, and initiator element (Inr) sites at or around your candidate transcription site in the genomic DNA sequence if it is available—you should usually be able to find either a TATA or an Inr.

FURTHER ANALYSIS AND USE OF RACE PRODUCTS

Cloning

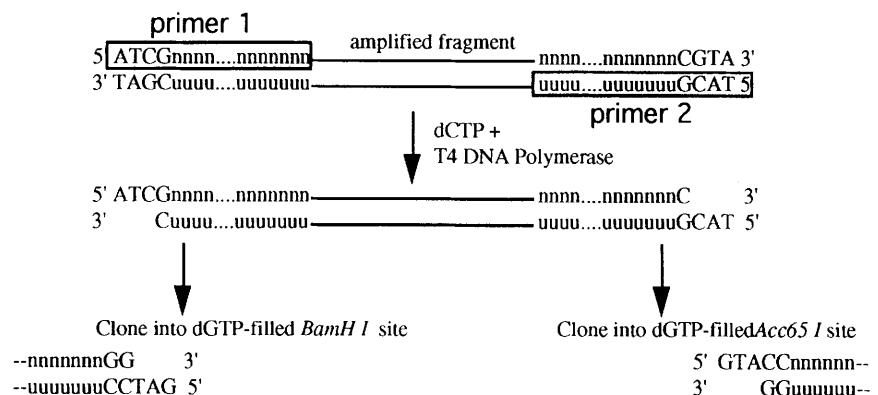
RACE products can be cloned like any other PCR products.

Option 1 To clone the cDNA ends directly from the amplification reaction (or after gel purification, which is recommended), ligate an aliquot of the

products to plasmid vector encoding a 1-nucleotide 3' overhang consisting of a T on both strands. Such vector DNA is available commercially (TA Kit, Invitrogen) or can be prepared easily and inexpensively.^(16,29–32)

Option 2 The classic RACE Q₁ primer encodes *Hind*III, *Sst*I, and *Xho*I restriction enzyme sites. Products can be cloned efficiently into vectors that have been double-cut with one of these enzymes and with a blunt-cutting enzyme such as *Sma*I (Note: Remember to “polish” the amplification products with Klenow enzyme or T4 DNA polymerase and separate them from residual *Taq* polymerase and dNTPs before carrying out the restriction enzyme digest). If clones are not obtained, determine whether the restriction enzyme chosen is cutting the amplified gene fragment a second time, at some internal location in the new and unknown sequence. A somewhat easier strategy is to append a restriction site (**not** *Hind*III, *Sst*I, or *Xho*I) onto the 5' end of the GSP2 primer to allow for the creation of overhanging strands at both ends of the amplified product.

Option 3 A safer and very effective approach is to modify the ends of the primers to allow the creation of overhanging ends using T4 DNA polymerase to chew back a few nucleotides from the amplified product in a controlled manner and Klenow enzyme (or Sequenase) to partially fill in restriction enzyme-digested overhanging ends on the vector, as shown below (adapted from Stoker and Iwahana et al.).^(33,34) For another conceptual variation, see Rashtchian et al.⁽¹⁰⁾



Note: *Acc*65I recognizes the same sequence that *Kpn*I does, but creates a 5' instead of a 3' overhang.

The advantages of this approach are that (1) it eliminates the possibility that the restriction enzymes chosen for the cloning step will cleave the cDNA end in the unknown region, (2) vector dephosphorylation is not required because vector self-ligation is no longer possible, (3) which means that insert kinasing (and polishing) is not necessary, (4) which means that insert multimerization and fusion clones are not observed either. In addition, the procedure is more reliable than TA cloning.

Procedure

Insert Preparation Select a pair of restriction enzymes for which you can synthesize half-sites appended to PCR primers that can be chewed back to form the appropriate overhangs, as shown for *Bam*HI and *Acc*65I, above. *Not*I and *Eag*I also are compatible with *Bam*HI and *Acc*65I. For RACE cloning, add

ATCG to the 5' end of Q₁ or NRC-2, and add TACG to the 5' end of GSP2. Carry out PCR as usual.

After PCR, add proteinase K (10 mg/ml stock) to the PCR reaction to a concentration of 50 µg/ml and incubate at 37°C for 30 min, to remove sticky *Taq* polymerase from the amplified DNA.⁽³⁵⁾ Extract the PCR products with phenol/CHCl₃, then CHCl₃ (but do not precipitate), to remove proteins. Filter the PCR products through a Microcon 100 spin column (or Microcon 30 if your product is <150 bp) three times using TE (not water) as the wash buffer, to remove unwanted organics, primers, and dNTPs.

Then, on ice, add the selected dNTP(s) (e.g., dCTP) to a final concentration of 0.2 mM, 1/10 volume of 10× T4 DNA polymerase buffer, and 1–2 units of T4 DNA polymerase. Incubate at 12°C for 15 min, and then at 75°C for 10 min to heat inactivate the T4 DNA polymerase. (Optional: Gel isolate DNA fragment of interest, depending on degree of success of PCR amplification.)

Vector Preparation Digest vector (e.g., pBluescript, Stratagene) using the selected enzymes (e.g., *Acc65I* and *BamHI*) under optimal conditions, in a volume of 10 µl. Then, add a 10-µl mixture containing the selected dNTP(s) (e.g., dGTP) at a final concentration of 0.4 mM, 1 µl of the restriction buffer used for digestion, 0.5 µl of Klenow, and 0.25 µl of Sequenase. Incubate at 37°C for 15 min, and then 75°C for 10 min to heat inactivate the polymerases. Gel-isolate the linearized vector fragment. For ligation, use equal molar amounts of vector and insert.

Sequencing

RACE products can be sequenced directly on a population level using a variety of protocols, including cycle sequencing, from the end at which the gene-specific primers are located. Note that the classic RACE products can not be sequenced on a population level using the Q₁ primer at the unknown end, because individual cDNAs contain different numbers of A residues in their poly(A) tails and as a consequence, the sequencing ladder falls out of register after reading through the tail. The 3'-end products can be sequenced from their unknown end using the following set of primers (TTTTTTTTTTTTTTTTTA, TTTTTTTTTTTTTTTTG, TTTTTTTTTTTTTTTTC). The non-T nucleotide at the 3' end of the primer forces the appropriate primer to bind to the inner end of the poly(A) tail.⁽³⁶⁾ The other two primers do not participate in the sequencing reaction. Individual cDNA ends, once cloned into a plasmid vector, can be sequenced from either end using gene-specific or vector primers.

Hybridization Probes RACE products are generally pure enough so that they can be used as probes for RNA and DNA blot analyses. It should be kept in mind that small amounts of contaminating nonspecific cDNAs will always be present. It is also possible to include a T7 RNA polymerase promoter in one or both primer sequences and to use the RACE products in *in vitro* transcription reactions to produce RNA probes.⁽⁵⁾ Primers encoding the T7 RNA polymerase promoter sequence do not appear to function as amplification primers as efficiently as those listed in the legend to Figure 1 (M. Frohman, pers. observation). Thus, the T7 RNA polymerase promoter sequence should not be incorporated into RACE primers as a general rule.

Construction of Full-length cDNAs It is possible to use the RACE protocol to create overlapping 5' and 3' cDNA ends that can later, through judicious choice of restriction enzyme sites, be joined together through subcloning to form a full-length cDNA. It is also possible to use the sequence information gained from acquisition of the 5' and 3' cDNA ends to make new primers representing the extreme 5' and 3' ends of the cDNA, and to employ them to

amplify a de novo copy of a full-length cDNA directly from the 3'-end cDNA pool. Despite the added expense of making two more primers, there are several reasons why the second approach is preferred.

First, a relatively high error rate is associated with the PCR conditions for which efficient RACE amplification takes place, and numerous clones may have to be sequenced to identify one without mutations. In contrast, two specific primers from the extreme ends of the cDNA can be used under inefficient but low-error rate conditions⁽³⁷⁾ for a minimum of cycles to amplify a new cDNA that is likely to be free of mutations. Second, convenient restriction sites are often not available, thus making the subcloning project difficult. Third, by using the second approach, the synthetic poly(A) tail (if present) can be removed from the 5' end of the cDNA. Homopolymer tails appended to the 5' ends of cDNAs have, in some cases, been reported to inhibit translation. Finally, if alternate promoters, splicing, and polyadenylation signal sequences are being used and result in multiple 5' and 3' ends, it is possible that one might join two cDNA halves that are never actually found together in vivo. Employing primers from the extreme ends of the cDNA as described confirms that the resulting amplified cDNA represents an mRNA actually present in the starting population.

TROUBLESHOOTING AND CONTROLS

Problems With Reverse Transcription and Prior Steps

Damaged RNA Electrophorese RNA in 1% formaldehyde minigel and examine integrity of the 18S and 28S ribosomal bands. Discard the RNA preparation if ribosomal bands are not sharp.

Contaminants Ensure that the RNA preparation is free of agents that inhibit reverse transcription, for example, LiCl and SDS (see Sambrook et al. regarding the optimization of reverse transcription reactions).⁽³⁸⁾

Bad Reagents To monitor reverse transcription of the RNA, add 20 μ Ci of [³²P]dCTP to the reaction, separate newly created cDNAs using gel electrophoresis, wrap the gel in Saran Wrap, and expose it to X-ray film. Accurate estimates of cDNA size can be best determined using alkaline agarose gels, but a simple 1% agarose minigel will suffice to confirm that reverse transcription took place and that cDNAs of reasonable length were generated. Note that adding [³²P]dCTP to the reverse transcription reaction results in the detection of cDNAs synthesized both through the specific priming of mRNA and through RNA self-priming. When a gene-specific primer is used to prime transcription (5' end RACE) or when total RNA is used as a template, the majority of the labeled cDNA will actually have been generated from RNA self-priming. To monitor extension of the primer used for reverse transcription, label the primer using T4 DNA kinase and [γ -³²P]ATP prior to reverse transcription. Much longer exposure times will be required to detect the labeled primer extension products than when [³²P]dCTP is added to the reaction.

To monitor reverse transcription of the gene of interest, one may attempt to amplify an internal fragment of the gene containing a region derived from two or more exons, if sufficient sequence information is available.

Problems With Tailing

Bad Reagents Tail 100 ng of a DNA fragment of ~100–300 bp for 30 min. In addition, mock-tail the same fragment (add everything but the TdT). Run

both samples in a 1% agarose minigel. The mock-tailed fragment should run as a tight band. The tailed fragment should have increased in size by 20–200 bp and should appear to run as a diffuse band that trails off into higher molecular weight products. If this not observed, replace reagents.

Mock-tail 25% of the cDNA pool (add everything but the TdT) Dilute to the same final concentration as the tailed cDNA pool. This serves two purposes. First, while amplification products will be observed using both tailed and untailed cDNA templates, the actual pattern of bands observed should be different. In general, discrete bands are observed using untailed templates after the first set of cycles, and a broad smear of amplified cDNA accompanied by some individual bands is typically observed using tailed templates. If the two samples appear different, this confirms that tailing took place and that the oligo(dT)-tailed Q_T primer is annealing effectively to the tailed cDNA during PCR. Second, observing specific products in the tailed amplification mixture that are not present in the untailed amplification mixture indicates that these products are being synthesized off the end of an A-tailed cDNA template, rather than by annealing of the dT-tailed primer to an A-rich sequence in or near the gene of interest.

Problems with Amplification

No Product If no products are observed for the first set of amplifications after 30 cycles, add fresh *Taq* polymerase and carry out an additional 15 rounds of amplification (extra enzyme is not necessary if the entire set of 45 cycles is carried out without interruption at cycle 30). Product is always observed after a total of 45 cycles if efficient amplification is taking place. If no product is observed, carry out a PCR reaction using control templates and primers to ensure the integrity of the reagents.

Smear Product From the Bottom of the Gel to the Loading Well Too many cycles, or too much starting material.

NonSpecific Amplification, But No Specific Amplification Check sequence of cDNA and primers. If all are correct, examine primers (using computer program) for secondary structure and self-annealing problems. Consider ordering new primers. Determine whether too much template is being added, or if the choice of annealing temperatures could be improved.

Alternatively, secondary structure in the template may be blocking amplification. Consider adding formamide⁽³⁹⁾ or ⁷aza-GTP (in a 1:3 ratio with dGTP) to the reaction to assist polymerization. ⁷aza-GTP can also be added to the reverse transcription reaction.

The Last Few Base Pairs of the 5'-End Sequence do not Match the Corresponding Genomic Sequence Be aware that reverse transcriptase and T7 and T3 RNA Polymerase can add on a few extra template-independent nucleotides.

Inappropriate Templates To determine whether the amplification products observed are being generated from cDNA or whether they derive from residual genomic DNA or contaminating plasmids, pretreat an aliquot of the RNA with RNase A.

CONCLUSIONS

The RACE protocol offers several advantages over conventional library screening to obtain additional sequence for cDNAs already partially cloned.

RACE is cheaper, much faster, requires very small amounts of primary material, and provides rapid feedback on the generation of the desired product. Information regarding alternate promoters, splicing, and polyadenylation signal sequences can be obtained and a judicious choice of primers (e.g., within an alternately spliced exon) can be used to amplify a subpopulation of cDNAs from a gene for which the transcription pattern is complex. Furthermore, differentially spliced or initiated transcripts can be separated by electrophoresis and cloned separately, and essentially unlimited numbers of independent clones can be generated to examine rare events. Finally, for 5' end amplification, the ability of reverse transcriptase to extend cDNAs all the way to the ends of the mRNAs is greatly increased, as a primer extension library is created instead of a general-purpose one.

ACKNOWLEDGMENTS

The work presented here was supported by a grant from the National Institutes of Health (HD29758).

REFERENCES

1. Frohman, M.A., M.K. Dush, and G.R. Martin. 1988. Rapid production of full-length cDNAs from rare transcripts by amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci.* **85**: 8998–9002.
2. Loh, E.L., J.F. Elliott, S. Cwirla, L.L. Lanier, and M.M. Davis. 1989. Polymerase chain reaction with single sided specificity: Analysis of T cell receptor delta chain. *Science* **243**: 217–220.
3. Ohara, O., R.I. Dorit, and W. Gilbert. 1989. One-sided PCR: The amplification of cDNA. *Proc. Natl. Acad. Sci.* **86**: 5673–5677.
4. Frohman, M.A. 1989. Creating full-length cDNAs from small fragments of genes: Amplification of rare transcripts using a single gene-specific oligonucleotide primer. In *PCR protocols and applications: A laboratory manual* (ed. M. Innis, D. Gelfand, J. Sninsky, and .. White), pp. 28–38. Academic Press, New York.
5. Frohman, M.A. and G.R. Martin. 1989. Rapid amplification of cDNA ends using nested primers. *Techniques* **1**: 165–173.
6. Dumas, J.B., M. Edwards, J. Delort, and J. Mallet. 1991. Oligodeoxyribonucleotide ligation to single-stranded cDNAs: A new tool for cloning 5' ends of mRNAs and for constructing cDNA libraries by *in vitro* amplification. *Nucleic Acids Res.* **19**: 5227–5223.
7. Fritz, J.D., M.L. Greaser, and J.A. Wolff. 1991. A novel 3' extension technique using random primers in RNA-PCR. *Nucleic Acids Res.* **19**: 3747.
8. Borson, N.D., W.L. Salo, and L.R. Drewes. 1992. A lock-docking oligo(dT) primer for 5' and 3' RACE PCR. *PCR Methods Applic.* **2**: 144–148.
9. Jain, R., R.H. Gomer, and J.J.J. Murtagh. 1992. Increasing specificity from the PCR-RACE technique. *BioTechniques* **12**: 58–59.
10. Rashtchian, A., G.W. Buchman, D.M. Schuster, and M.S. Berninger. 1992. Uracil DNA glycosylase-mediated cloning of PCR-amplified DNA: Application to genomic and cDNA cloning. *Anal. Biochem.* **206**: 91–97.
11. Schuster, D.M., G.W. Buchman, and A. Rashtchian. 1992. A simple and efficient method for amplification of cDNA ends using 5' RACE. *Focus* **14**: 46–52.
12. Bertling, W.M., F. Beier, and E. Reichenberger. 1993. Determination of 5' ends of specific mRNAs by DNA ligase-dependent amplification. *PCR Methods Applic.* **3**: 95–99.
13. Frohman, M.A. 1993. Rapid amplification of cDNA for generation of full-length cDNA ends: Thermal RACE. *Methods Enzymol.* **218**: 340–356.
14. Monstein, H.J., J.U. Thorup, R. Folkesson, A.H. Johnsen, and J.F. Rehfeld. 1993. cDNA deduced procionin—structure and expression in protochordates resemble that of procholecystokinin in mammals. *FEBS Lett.* **331**: 60–64.
15. Templeton, N.S., E. Urcelay, and B. Safer. 1993. Reducing artifact and increasing the yield of specific DNA target fragments during PCR-RACE or anchor PCR. *BioTechniques* **15**: 48–50.
16. Frohman, M.A. 1994. Cloning PCR products: Strategies and tactics. In *PCR: The polymerase chain reaction. Methods in molecular biology series* (ed. K.B. Mullis, F. Ferre, R.A. Gibbs), pp. 14–37. Birkhäuser, Boston, MA.
17. Tessier, D.C., R. Brousseau, and T. Vernet. 1986. Ligation of single-stranded oligodeoxyribonucleotides by T4 RNA ligase. *Anal. Biochem.* **158**: 171–178.
18. Mandl, C.W., F.X. Heinz, E. Puchhammer-Stockl, and C. Kunz. 1991. Sequencing the termini of capped viral RNA by 5'-3' ligation and PCR. *BioTechniques* **10**: 484–486.
19. Volloch, V., B. Schweizer, X. Zhang, and S. Rits. 1991. Identification of negative-strand

- complements to cytochrome oxidase subunit III RNA in *Trypanosoma brucei*. *Biochemistry* **88**: 10671–10675.
20. Brock, K.V., R. Deng, and S.M. Riblet. 1992. Nucleotide sequencing of 5' and 3' termini of bovine viral diarrhea virus by RNA ligation and PCR. *Viol. Methods* **38**: 39–46.
 21. Bertrand, E., M. Fromont-Racine, R. Pictet, and T. Grange. 1993. Visualization of the interaction of a regulatory protein with RNA *in vivo*. *Proc. Natl. Acad. Sci.* **90**: 3496–3500.
 22. Fromont-Racine, M., E. Bertrand, R. Pictet, and T. Grange. 1993. A highly sensitive method for mapping the 5' termini of mRNAs. *Nucleic Acids Res.* **21**: 1683–1684.
 23. Liu, X. and M.A. Gorovskiy. 1993. Mapping the 5' and 3' ends of tetrahymena-thermophila mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). *Nucleic Acids Res.* **21**: 4954–4960.
 24. Sallie, R. 1993. Characterization of the extreme 5' ends of RNA molecules by RNA ligation-PCR. *PCR Methods Applic.* **3**: 54–56.
 25. Wickens, M. and P. Stephenson. 1984. Role of the conserved AAUAAA sequence: Four AAUAAA point mutants prevent mRNA 3' end formation. *Science* **226**: 1045–1050.
 26. Coleclough, C. 1987. Use of primer-restriction end adapters in cDNA cloning. *Methods Enzymol.* **154**: 64–83.
 27. Frohman, M.A., M.E. Dickinson, B.L.M. Hogan, and G.R. Martin. 1993. Localization of two new and related homeobox-containing genes to chromosomes 1 and 5, near the phenotypically similar mutant loci *dominant hemimelia (Dh)* and *hemimelic extra-toes (Hx)*. *Mouse Genome* **91**: 323–325.
 28. Don, R.H., P.T. Cox, B.J. Wainwright, K. Baker, and J.S. Mattick. 1991. Touchdown PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Res.* **19**: 4008.
 29. Mead, D.A., N.K. Pey, C. Herrnstadt, R.A. Marcil, and L.A. Smith. 1991. A universal method for direct cloning of PCR amplified nucleic acid. *Biotechnology* **9**: 657–663.
 30. Marchuk, D., M. Drumm, A. Saulino, and F.S. Collins. 1991. Construction of T-vector, a rapid and general system for direct cloning of unmodified PCR products. *Nucleic Acids Res.* **19**: 1154.
 31. Kovalic, D., J.H. Kwak, and B. Weisblum. 1991. General method for direct cloning of DNA fragments generated by the polymerase chain reaction. *Nucleic Acids Res.* **19**: 4650.
 32. Holton, T.A. and M.W. Graham. 1991. A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucleic Acids Res.* **19**: 1156.
 33. Stoker, A.W. 1990. Cloning of PCR products after defined cohesive termini are created with T4 DNA polymerase. *Nucleic Acids Res.* **18**: 4290.
 34. Iwahana, H., N. Mizusawa, S. Ii, K. Yoshimoto, and M. Itakura. 1994. An end-trimming method to amplify adjacent cDNA fragments by PCR. *BioTechniques* **16**: 94–98.
 35. Crowe, J.S., H.J. Cooper, M.A. Smith, M.J. Sims, D. Parker, and D. Gewert. 1991. Improved cloning efficiency of polymerase chain reaction (PCR) products after proteinase K digestion. *Nucleic Acids Res.* **19**: 184.
 36. Thweatt, R., S. Goldstein, and R.J.S. Reis. 1990. A universal primer mixture for sequence determination at the 3' ends of cDNAs. *Anal. Biochem.* **190**: 314.
 37. Eckert, K.A. and T.A. Kunkel. 1990. High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res.* **18**: 3739–3745.
 38. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
 39. Sarker, G., S. Kapelner, and S.S. Sommer. 1990. Formamide can dramatically improve the specificity of PCR. *Nucleic Acids Res.* **18**: 7465.