One-sided polymerase chain reaction: The amplification of cDNA

[direct sequencing/tropomyosin/zebrafish (Brachydanio rerio)/European common frog (Rana temporaria)/sequence evolution]

OSAMU OHARA*, ROBERT L. DORIT†, AND WALTER GILBERT

Department of Cellular and Developmental Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138

Contributed by Walter Gilbert, March 29, 1989

ABSTRACT   We report a rapid technique, based on the polymerase chain reaction (PCR), for the direct targeting, enhancement, and sequencing of previously uncharacterized cDNAs. This method is not limited to previously sequenced transcripts, since it requires only two adjacent or partially overlapping specific primers from only one side of the region to be amplified. These primers can be located anywhere within the message. The specific primers are used in conjunction with nonspecific primers targeted either to the poly(A)$^+$ region of the message or to an enzymatically synthesized d(A) tail. Pairwise combinations of specific and general primers allow for the amplification of regions both 3' and 5' to the point of entry into the message. The amplified PCR products can be cloned, sequenced directly by genomic sequencing, or labeled for sequencing by amplifying with a radioactive primer. We illustrate the power of this approach by deriving the cDNA sequences for the skeletal muscle α-tropomyosins of European common frog (Rana temporaria) and zebrafish (Brachydanio rerio) using only 300 ng of a total poly(A)$^+$ preparation. In these examples, we gained initial entry into the tropomyosin messages by using heterologous primers (to conserved regions) derived from the rat skeletal muscle α-tropomyosin sequence. The frog and zebrafish sequences are used in an analysis of tropomyosin evolution across the vertebrate phylogenetic spectrum. The results underscore the conservative nature of the tropomyosin molecule and support the notion of a constrained heptapeptide unit as the fundamental structural motif of tropomyosin.

The polymerase chain reaction (PCR) is a powerful method for the enrichment of specific target sequences. By carrying out repeated cycles of annealing, synthesis, and denaturation in the presence of a thermostable DNA polymerase [Thermus aquaticus DNA polymerase (Taq polymerase)], sequences bounded by a pair of short, sequence-specific primers can be amplified more than a millionfold (1). However, the dependence of the PCR technique on a pair of specific primers flanking the region of interest has generally limited its use to targets whose sequence is known a priori.

Frequently, one is interested in sequences that have not been as yet fully characterized. A method requiring only a single small region of known sequence (from which specific primers can be designed) while still taking advantage of the power of PCR would substantially facilitate the extraction and sequencing of DNA molecules.

We report here a general method for the amplification of previously unsequenced cDNA target molecules. Fig. 1a illustrates schematically the three-pronged approach we have developed. First, we gain access into a transcript by isolating and amplifying a core region, using two primers derived from related sequences. Then, based on this core sequence, we design primers that, in combination with a second nonspecific primer [complementary either to the 3' poly(A) tail or to an

enzymatically synthesized tail at the 5' end], permit the amplification of the regions both upstream and downstream of the core sequence. The combination of a single specific primer and a nonspecific primer generally is not sufficient to yield a unique product. Consequently, we reamplify this mix using a second specific primer, adjacent to the first, and the same nonspecific primer. This results in the amplification of a unique sequence.

These amplified fragments can be cloned into vectors from which the insert can be subsequently sequenced (2); alternatively, they can be directly sequenced by using either the "genomic sequencing" (3) or the dideoxy chain-termination method (4). We have also used a rapid, direct, chemical sequencing strategy that capitalizes on the specificity of the PCR method. A single primer is labeled by using any of the currently available tagging methods (see Materials and Methods) and then used to reamplify a sample previously enriched for a particular sequence. This produces double-stranded target molecules, labeled exclusively at one of the 5' ends, that then can be chemically sequenced (5, 6) and visualized directly on a sequencing gel.

To illustrate the power of these techniques, we extracted (in parallel) and sequenced the complete transcripts of skeletal muscle α-tropomyosin from both the European common frog (Rana temporaria) and the zebrafish (Brachydanio rerio).

Other methods that seek to extend the use of PCR to previously unknown sequences have recently been reported. The restriction digestion and subsequent circularization of genomic DNA (inverse polymerase chain reaction, IPCR) permits the amplification of unique sequences (7). Methods involving the tailing of cDNA [the "RACE" (8) and "A-PCR" (9) approaches] permit the subsequent cloning of uncharacterized or variable messages.

MATERIALS AND METHODS

Preparation of cDNA. Poly(A)$^+$ RNA was prepared from frog leg muscle and from complete zebrafish according to standard methods (10). Poly(A)$^+$ RNA was primed by using an oligo(dT) 20-mer. cDNA synthesis on 100 ng of poly(A)$^+$ RNA was accomplished by using Moloney murine leukemia virus (Mo-MLV) reverse transcriptase in a volume of 5 μl (11).

Amplification of Core Region. Following cDNA synthesis as described above, 100 pmol of each of a pair of specific (imperfect) oligomers was added, in water, bringing the volume to 75 μl. The mixture was boiled for 3 min and quenched on ice. Remaining PCR reagents were then added. Amplifications were done in 100 μl of 50 mM KCl, 10 mM Tris·HCl (pH 8.8), 1.5 mM MgCl$_2$, 3 mM dithiothreitol, 0.1 mg of bovine serum albumin per ml, 200 μM (each) dNTPs, and

Abbreviations: PCR, polymerase chain reaction; Mo-MLV, Moloney murine leukemia virus.
*Present address: Shionogi Research Laboratories, Fukushima-ku, Osaka 553, Japan.
†To whom reprint requests should be addressed.

2.5 units of Taq polymerase. Thirty cycles of PCR were carried out: annealing at 45°C, 2 min; extension at 72°C, 3 min; denaturation at 94°C, 1 min; final extension, 5 min.

Amplification of 3' Region. The amplifications for the 3' region were carried out in a final 100-μl volume, using a 10-μl sample of the above "core amplification" mixture, to which 100 pmol of both a nested specific primer and the nonspecific (dT)$_{20}$ primer were added. Remaining reagents and PCR conditions are as above.

Synthesis of Specific cDNA Molecules for Amplification of 5' Region. A specific oligomer was used to prime the Mo-MLV reverse transcriptase on the poly(A)$^+$ RNA. Initial annealing of the oligomer (1 fmol) was done in 5 μl containing 0.2 M NaCl, 40 mM Tris·HCl (pH 7.5), 5 mM EDTA, and 100 ng of poly(A)$^+$ RNA, incubated at 65°C for 5 min, placed at 40°C for 4 hr, ethanol precipitated, resuspended in 25 μl of reverse transcriptase mix containing 50 mM Tris·HCl (pH 8.3), 75 mM KCl, 10 mM dithiothreitol, 3 mM MgCl$_2$, 50 μg of Actinomycin D per ml, 0.5 mM (each) dNTPs, and 200 units of Mo-MLV reverse transcriptase (BRL), and incubated at 37°C for 1 hr.

Poly[d(A)] Tailing of Specific cDNA Molecules and Amplification of 5' Region. Following the synthesis described above, the mixture was phenol extracted, ethanol precipitated twice, dried under vacuum, resuspended in 6.5 μl of water, boiled for 2 min, quenched on ice, and tailed in a final 10-μl volume using terminal deoxynucleotidyltransferase, according to standard protocols (12). The products were ethanol precipitated and resuspended in 100 μl of the PCR mixture as described in "Amplification of Core Region." PCR conditions were as follows: 40 cycles, annealing at 50°C, 30 sec; extension at 72°C, 1 min; denaturation at 94°C for 40 sec, in the presence of 100 pmol of the specific oligomer and (dT)$_{20}$. A 10-μl sample of the above was reamplified in 100-μl final volume with a nested, specific oligomer, internal to the

first. PCR conditions were as follows: 55°C, 30 sec; 72°C, 1 min; 94°C, 40 sec; 30 cycles.

Direct Sequencing of PCR Products. Ten picomoles of a specific PCR primer was labeled with T4 polynucleotide kinase (NEB) and 200 μCi of [γ-^{32}P]ATP (6000 Ci/mmol; 1 Ci = 37 GBq) under conditions specified by the supplier (37°C for 30 min). The 10-μl kinase mixture was placed directly into 90 μl of standard PCR mixture containing 100 pmol of unlabeled second primer and 1 ng of template DNA. Five to 10 PCR cycles generate enough labeled DNA for direct chemical sequencing. Labeled fragment was gel purified in a 5% native acrylamide gel and recovered by absorbing onto a DEAE membrane (13).

RESULTS

The results of our three-pronged approach (Fig. 1a) are described below.

Sequencing the Core Region: Use of Imperfect Primers Derived from Homologous Sequences. To enter the sequence of the unknown tropomyosin cDNAs, we first identified regions of high amino acid sequence conservation in previously sequenced tropomyosin genes (14–18). We chose primers complementary to those conserved stretches that contained amino acids with low codon degeneracy and synthesized two sets of "21-mers," R1/R2 and R3/R4, derived from the rat tropomyosin nucleotide sequence (15, 16). Fig. 1b depicts the positions of all primers used in this study, and Table 1 collects the sequences of all the oligonucleotides described in this paper.

We prepared poly(A)$^+$ RNA from frog leg muscle and from a complete zebrafish homogenate. Tropomyosin transcripts appear to comprise >0.1% of the total muscle poly(A)$^+$ RNA pool (19). A complete array of cDNA molecules was then generated from the poly(A)$^+$ RNA mixture using Mo-MLV
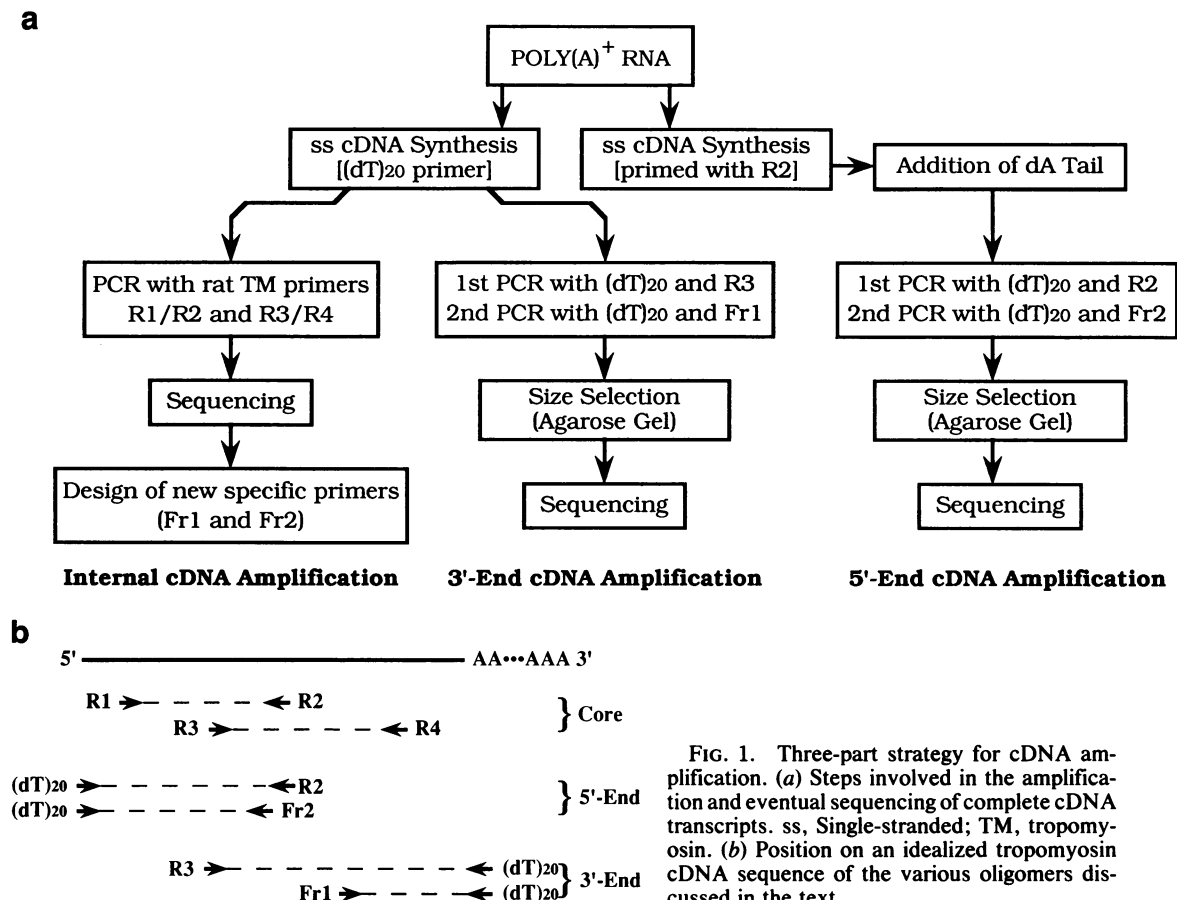

FIG. 1. Three-part strategy for cDNA amplification. (a) Steps involved in the amplification and eventual sequencing of complete cDNA transcripts. ss, Single-stranded; TM, tropomyosin. (b) Position on an idealized tropomyosin cDNA sequence of the various oligomers discussed in the text.

Table 1. Oligonucleotide primer sequences

	PCR primer
R1:	5'-ATGGACGCCATCAAGAAGAAG-3' (1)
R2:	5'-CTCCATCTTCTCCTCATCCTT-3' (425)
R3:	5'-ATCCAGCTGGTTGAGGAGGAG-3' (273)
R4:	5'-GGTCTCAGCCTCCTTCAGCTT-3' (710)
FR1:	5'-GAGCTTGAGGAAGAGTTGAAAA-3' (573)
FR2:	5'-GTTCTCAATGACTTTCATGCCTC-3' (395)
	Sequencing/probing primer
FR3:	5'-CTGTTCTGCTCTGTCCAAGG-3' (71)
FR4:	5'-TATGCCCAGAAACTGAAGTA-3' (780)

The prefixes indicate the source of the sequence information (R = rat; FR = frog). Numbers in parentheses at the right indicate nucleotide position on the target sequence where the 5' end of the oligomer will anneal. Position 1 is the first nucleotide in the initiation methionine.

reverse transcriptase (11) primed with an oligo(dT) 20-mer. The single-stranded cDNA resulting from this synthesis became the template for PCR amplifications using the two sets of imperfect primers. As shown in Fig. 1*b*, these imperfect primer pairs each spanned ≈400 nucleotides, with a 100-nucleotide region of overlap. Thus we sought to amplify ≈700 nucleotides of the tropomyosin transcript core region.

The results of the amplifications on frog and zebrafish cDNA are shown in the left panel (lanes 1–4) of Fig. 2*a*. Although a major discrete band of the expected size appears, there might have been two or more isotypes of tropomyosin present in the original cDNA preparations, and the amplified product might have contained a mixture of different tropomyosin transcripts. To ensure that a homogeneous transcript was sequenced, the amplified products were cut out of the 2% agarose gel, eluted, and cloned into the "Bluescript" plasmid vector (Stratagene) using the G-C tailing method (12). We sequenced six colonies (20, 21) containing frog muscle cDNA amplified with the R3/R4 primer pair: all were identical and contained only the skeletal muscle α transcript. In contrast, five colonies derived from the amplification of zebrafish cDNA with the R3/R4 primer pair proved heterogeneous upon sequencing: three colonies carried skeletal muscle α transcripts, one carried smooth muscle β transcript, and one contained the skeletal muscle β transcript. Since our focus was exclusively on the major isotype of tropomyosin, and we favor methods that do not involve the cloning and sequencing

of a single amplified molecule, given concerns about the fidelity of Taq polymerase (22), we used direct sequencing of the PCR product as frequently as possible. We obtained the complete sequence of the core region directly from the amplified fragments by using a combination of genomic sequencing and the direct PCR-mediated sequencing method. Even in those cases in which a mixture of templates was present, the direct method allowed us to read the sequence of the major cDNA isotype. Fig. 3 shows a representative fragment of a sequencing gel obtained by this method. Approximately 700 nucleotides of sequence were thus generated from each of the two tropomyosin cDNA sequences. Nonetheless, substantial portions of the transcript, both upstream and downstream of the amplified fragments, remained to be explored.

Characterizing the 3' Region: Enrichment Amplification Using One-Sided Specific Primers. To extend our analysis into the 3' region of the transcript, we amplified the cDNA mixture prepared for the previous step, using equimolar amounts of the single specific primer R3 (originally thought to be imperfectly matched) located in the previously derived core region, and the $(dT)_{20}$ primer. After 30 cycles of amplification, a large number of different sequences appeared, as shown by the smears in lanes 5 and 8 of Fig. 2*a*. This amplified mixture nevertheless did include the sequence of interest— the 3' region of the tropomyosin transcript—since we could detect it on a Southern blot by probing with an internal oligomer (Fr4) (lanes 5 and 8, Fig. 2*b*).

To rescue the appropriate sequence, a 10% aliquot of the above amplification was reamplified directly, using 100 pmol of a new, nested specific oligomer (Fr1) and an equimolar amount of the $(dT)_{20}$ oligomer. Lanes 6 and 9 of Fig. 2*a* (and lanes 6 and 9, Fig. 2*b*) show that there is still a heterogeneous mixture of amplified fragments, most likely reflecting the distribution of $poly(A)^+$ tail lengths. The uppermost band on the 2% agarose gel (≈400–500 nucleotides long), which appeared to contain our target molecule, was cut out, eluted, and reamplified using the same pair of primers [$(dT)_{20}$ and Fr1]. This size-selection reduced the potential targets for amplification and increased the specificity of subsequent PCR rounds. Lanes 7 and 10, Fig. 2*a*, show that this resulted in a single band, which contained the desired sequence (as seen in lanes 7 and 10, Fig. 2*b*, probed with oligo Fr4). We sequenced the amplified product using both the genomic
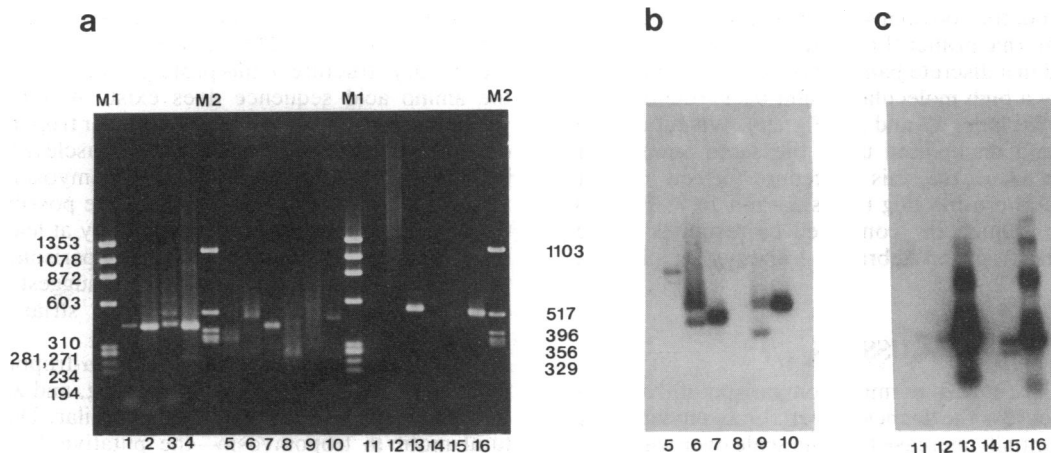

Fig. 2. (*a*) Agarose gels stained with ethidium bromide showing the results of amplifications of core, 3', and 5' regions. Lanes M1 and M2, DNA size markers (given in nucleotides) (ϕX *Hae* III digest and Bluescript *Hin*fI digest). Lanes 1–4, results of core region amplification: frog cDNA, primers R1/R2 (lane 1); primers R3/R4 (lane 2); fish cDNA, primers R1/R2 (lane 3); primers R3/R4 (lane 4). Lanes 5–10, 3' end amplifications: frog cDNA, primers $(dT)_{20}$/R3 (lane 5); primers $(dT)_{20}$/Fr1 (lane 6); size-selected frog amplification, primers $(dT)_{20}$/Fr1 (lane 7); fish cDNA, primers $(dT)_{20}$/R3 (lane 8); primers $(dT)_{20}$/Fr1 (lane 9); size-selected fish amplification, primers $(dT)_{20}$/Fr1 (lane 10). Lanes 11–16, 5' end amplifications: frog cDNA, primers $(dT)_{20}$/R2 (lane 11); primers $(dT)_{20}$/Fr2 (lane 12); size-selected frog amplification, primers $(dT)_{20}$/Fr2 (lane 13); fish cDNA, primers $(dT)_{20}$/R2 (lane 14); primers $(dT)_{20}$/Fr2 (lane 15); size-selected fish amplification, primers $(dT)_{20}$/Fr2 (lane 16). (*b*) Southern blot of lanes 5–10, probed with oligomer Fr4. (*c*) Southern blot of lanes 11–16, probed with oligomer Fr3.
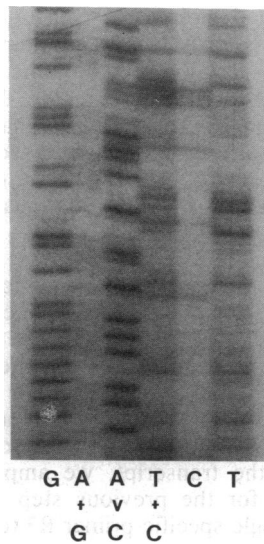
FIG. 3. Fragment of sequencing gel obtained by direct PCR-mediated method. The fragment shown corresponds to the internal core region of zebrafish tropomyosin cDNA (nucleotide positions 340–400). Sequencing reactions are as described in refs. 5 and 6.

Table 2. Summary of zebrafish (*B. rerio*) and European common frog (*R. temporaria*) tropomyosin cDNA sequences and comparisons

	5' untranslated region	Coding region	3' untranslated region
B. rerio	87 nt	855 nt (284 aa)	75 nt
R. temporaria	109 nt	855 nt (284 aa)	72 nt
No. of nucleotide differences	*	130 nt Silent: 101 nt Replace: 29 nt (18 aa)	54 nt[†]
% difference		Silent: 35 Replace: 3.2[‡]	75.0

nt, Nucleotides; aa, amino acids.
*The 5' untranslated region shows two regions of sequence conservation [positions -109 to -74 (frog), -87 to -52 (fish), and -6 to 1 (frog and fish)].
[†]The 3' regions have been aligned without gaps.
[‡]Percentage differences have been calculated according to ref. 23. Shown are the weighted uncorrected averages. Corrected weighted averages for silent positions = 110.5%; for replacement positions = 3.8%.

sequencing and the direct sequencing methods previously described. In this way, we generated an additional 300 nucleotides [including about 100 residues of the poly(A) tail] in the case of the European common frog cDNA and ≈350 nucleotides of zebrafish tropomyosin sequence.

Characterizing the 5' Region: Single Primer Extension, Tailing, and Amplification. By using a specific oligomer (R2) complementary to the rat tropomyosin sequence, we primed the Mo-MLV reverse transcriptase on the poly(A)$^+$ RNA to generate cDNA molecules containing the 5' region of the transcript. We added poly(dA) tails using terminal deoxynucleotidyltransferase (12) and amplified these sequences by PCR, using the same specific primer (R2) and a (dT)$_{20}$ nonspecific primer complementary to the newly added tails. After 40 cycles of amplification, the PCR product could generally not be visualized with ethidium bromide staining (lanes 11 and 14, Fig. 2a) but did appear as a faint band on subsequent Southern blots probed with an internal probe (Fr3) (lanes 11 and 14, Fig. 2c; not visible in the reproduction). An aliquot of the above mixture, when reamplified with a new, nested internal primer (Fr2) and the nonspecific (dT)$_{20}$ primer, resulted in a discrete band. This band, however, was accompanied by a high molecular weight background (lanes 12 and 15, Fig. 2a; lanes 12 and 15, Fig. 2c). We cut out the discrete band and reamplified using the same primer pair [Fr2/(dT)$_{20}$]. In each case, this procedure yielded a major band, which Southern blotting (lanes 13 and 16 in Fig. 2c) and subsequent sequencing confirmed corresponds to the 5' regions of the frog and zebrafish tropomyosin cDNAs, respectively.

DISCUSSION

Our two cDNA sequences permit a comparison of tropomyosin sequences across the deepest of vertebrate phylogenetic forks and substantially increase the time scale over which the evolution of tropomyosin can be surveyed. The comparison reveals an extremely stable and conservative structural protein. [The tropomyosin cDNA sequences have been deposited in GenBank (accession nos. M24634 and M24635). Table 2 summarizes some properties of the sequences.] Despite the broad phylogenetic distance (≈400 million years since the last common ancestor) separating *R. temporaria* from *B. rerio*, the two tropomyosin proteins differ at only 19 of 284 amino

acids. This degree of conservation is probably due to the structural role this protein plays, interacting with other proteins (actin, troponin) in the muscle contraction apparatus (24, 25).

A large number of different tropomyosins have been identified, from both muscle and cytoplasmic sources. Muscle tissue (smooth and striated) may contain both α and β subunits of tropomyosin; up to nine isoforms of a given tropomyosin may be present as a result of alternative splicing (14, 26). This heterogeneity of possible tropomyosin forms will be reflected in a mixture of cDNAs within a single organism. A procedure such as ours, which used whole organisms (in the case of zebrafish) or striated muscle (in the case of the frog) to prepare mRNA, makes it difficult *a priori* to determine which cDNA is being isolated. To determine the identity of the cDNA sequences, we compared them to all available vertebrate tropomyosin sequences: quail (*Coturnix coturnix*) skeletal α-tropomyosin (27), chicken (*Gallus gallus*) smooth muscle α-tropomyosin (28), rabbit (*Oryctolaqus cuniculus*) muscle α-tropomyosin (29), rat (*Rattus norvegicus*) skeletal α- and β-tropomyosins (15, 30), and human (*Homo sapiens*) skeletal muscle α-tropomyosin and nonmuscle tropomyosin (31). Despite the extensive conservation in the primary structure of this protein, interspecific variation in the amino acid sequence does exist. A number of these variable sites are diagnostic for particular tropomyosin types, occurring (for example) in all striated muscle α-tropomyosins but not in smooth muscle or β-tropomyosins. Of the 86 variable sites in our comparison, 32 are possibly "informative"—a residue at that site is shared by at least two but no more than $n - 1$ sequences in the comparison. The 12 sites that are informative for this case all suggest that the two cDNA sequences presented here are striated muscle α-tropomyosin messages.

Table 3 compares the striated-muscle α-tropomyosins from five vertebrates (human, rat, quail, frog, and zebrafish). All of these sequences are remarkably similar. Does the structural motif in tropomyosin—the putative NxxNAxB heptaresidue repeat that runs throughout the sequence associated with the double coiled-coil structure of the molecule (33)—constrain this variation? In the five proteins compared, there are 52 variable sites; of these, 29 are at "x" positions within the hepatapeptide motif, whereas 11 occur at "N" (hydrophobic) positions, 2 at "A" (acidic) positions, and 5 at "B" (basic) positions. This distribution is significantly different from the uniform random expectation ($P < 0.05$), with

Biochemistry: Ohara *et al.*

Proc. Natl. Acad. Sci. USA 86 (1989) 5677

Table 3. Amino acid sequence divergence of available skeletal muscle α-tropomyosins

	Zebrafish	Frog	Quail	Rat	Human
Zebrafish		20 (7.0)	19 (6.7)	22 (7.7)	20 (7.0)
Frog	10 (6.6)		13 (4.6)	16 (5.6)	18 (6.3)
Quail	12 (7.9)	6 (3.9)		12 (4.2)	19 (6.7)
Rat	14 (9.2)	10 (6.6)	7 (4.6)		25 (8.8)
Human	10 (6.6)	7 (4.6)	8 (5.3)	13 (8.6)	

Figures above the diagonal indicate amino acid differences over the entire molecule; figures in parentheses show the percentage divergence between sequences. Data below the diagonal reflect amino acid changes and percentage divergence after excluding the most variable and alternatively spliced exons [1, 2, 6, and 9 (32)]. Percentages below the diagonal are calculated by using exons 3, 4, 5, 7, and 8 (total amino acids: 152).

an excess of substitutions at "x" positions that accords with the prediction based on the heptapeptide motif: the second, third, and fifth positions are less stringently constrained. Nonetheless, the small amount of variation in these sequences reflects a strong purifying selection acting on practically every residue—in the overwhelming majority of positions even equivalent substitutions have been weeded out. The 52 variant amino acids in our comparison occur at 38 positions; 47 of these changes are equivalent substitutions.

Curiously, the pattern of values in Table 3 is not what one expects for a slowly evolving gene. The sequence divergences are not congruent with the phylogenetic tree based on the species involved. (For example, the divergence between rat and human tropomyosins is greater than the divergence between frog and human tropomyosins.) This suggests strongly that there are several genes for the skeletal muscle α-tropomyosins, which diverged from a common ancestral gene prior to the vertebrate colonization of the land. This possibility is further supported by the distribution of parallel substitutions at identical positions. At 10 variable positions the same amino acid change occurs in two sequences, and in no case does the similarity reflect common ancestry.

We believe this method for obtaining and characterizing RNA transcripts represents a significant improvement over previous approaches and holds great promise for the comparative sequencing of homologous transcripts from different species. Initial entry into the cDNA sequence can be achieved by using imperfect oligomers derived from related organisms. The need for specific primers from only one side of the region to be amplified expands the power of PCR, permitting the amplification of previously uncharacterized messages. This method does not require library construction and screening, laborious and often unpredictable steps in the isolation of complete transcripts, and, in contrast with previous approaches, requires a far smaller amount of starting poly(A)$^+$ RNA—only 300 ng of poly(A)$^+$ RNA was utilized in determining the full sequences of both frog and zebrafish tropomyosin cDNAs. Finally, the method lends itself to the processing in parallel of cDNAs derived from a variety of different sources.

1. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239,** 487–491.
2. Scharf, S. J., Holm, G. T. & Erlich, H. A. (1986) *Science* **233,** 1076–1078.
3. Church, G. M. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.
4. Gyllensten, U. B. & Erlich, H. A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7652–7656.
5. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 560–564.
6. Rubin, C. M. & Schmid, C. W. (1980) *Nucleic Acids Res.* **8,** 4613–4619.
7. Ochman, H., Gerber, A. S. & Hartl, D. L. (1988) *Genetics* **120,** 621–623.
8. Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8998–9002.
9. Loh, E. Y., Elliott, J. F., Cwirla, S., Lanier, L. L. & Davis, M. M. (1989) *Science* **243,** 217–220.
10. Ausebel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds. (1987) *Current Protocols in Molecular Biology* (Wiley, New York), Units 4.3 and 4.5.
11. D'Alessio, J. M. & Gerard, G. F. (1988) *Nucleic Acids Res.* **16,** 1999–2014.
12. Eschenfeldt, W. H., Puskas, R. S. & Berger, S. L. (1987) *Methods Enzymol.* **152,** 337–342.
13. Dretzen, G., Bellard, M., Sassone-Corsi, P. & Chambon, P. (1981) *Anal. Biochem.* **112,** 295–298.
14. Ruiz-Opazo, N. & Nadal-Ginard, N. (1987) *J. Biol. Chem.* **262,** 4755–4765.
15. Helfman, D. M., Cheley, S., Kuismanen, E., Finn, L. A. & Yamawaki-Kataoka, Y. (1986) *Mol. Cell. Biol.* **6,** 3582–3595.
16. Mak, A. S., Smillie, L. B. & Stewart, G. R. (1980) *J. Biol. Chem.* **255,** 3647–3655.
17. MacLeod, A. R., Houlker, C., Reinach, F. C., Smillie, L. B., Talbot, K., Modi, G. & Walsh, F. S. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 7835–7839.
18. MacLeod, A. R. (1982) *Eur. J. Biochem.* **126,** 293–297.
19. Helfman, D. M., Feramsico, J. R., Fiddes, J. C., Thomas, G. P. & Hughes, S. H. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 31–35.
20. Church, G. M. & Kieffer-Higgins, S. (1988) *Science* **240,** 185–188.
21. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7,** 1513–1523.
22. Wilson, A. C. & Paabo, S. (1988) *Nature (London)* **334,** 387–388.
23. Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. & Dodgson, J. (1980) *Cell* **20,** 555–566.
24. Smillie, L. B. (1975) *J. Mol. Biol.* **98,** (2), 281–291.
25. Pearlstone, J. R. & Smillie, L. B. (1982) *J. Biol. Chem.* **257,** 10587–10592.
26. Giometti, C. S. & Anderson, N. L. (1984) *J. Biol. Chem.* **259,** 14113–14120.
27. Hastings, K. E. M. & Emerson, C. P., Jr. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 1553–1557.
28. Helfman, D. M., Feramisco, J. R., Ricci, W. M., Hughes, S. H. (1984) *J. Biol. Chem.* **259,** 14136–14143.
29. Putney, S. D., Herlihy, W. C. & Schimmel, P. (1983) *Nature (London)* **302,** 718–723.
30. Yamawaki-Katakoa, Y. & Helfman, D. M. (1985) *J. Biol. Chem.* **260,** 14440–14445.
31. Clayton, L., Reinach, F. C., Chumbley, G. M. & MacLeod, A. R. (1988) *J. Mol. Biol.* **201,** 507–515.
32. Colote, S., Widada, J. S., Ferraz, C., Bonhomme, F., Marti, J. & Liautard, J. P. (1988) *J. Mol. Evol.* **27,** 228–235.
33. McLachlan, A. D. & Stewart, M. (1975) *J. Mol. Biol.* **98,** 293–304.