Marcos Malumbres, Ramon Mangues, Neus Ferrer, Suying Lu and Angel Pellicer
New York University Medical Center
New York, NY, USA

Analysis of the Size Distribution of First-Strand cDNA Molecules

BioTechniques 22:1119-1126 (June 1997)

ABSTRACT

We have developed a method to analyze the size distribution of the first-strand cDNA molecules corresponding to given mRNA species. First-strand molecules synthesized from cytoplasmic polyadenylated RNAs are separated by electrophoresis on an alkaline agarose gel, and a Southern blot hybridization is performed. As an example, we analyzed the first-strand molecules corresponding to the human c-myc mRNAs. This method can be used to determine whether full-length, first-strand molecules corresponding to an mRNA species to be cloned are synthesized efficiently. Interestingly, this method allows one to analyze full-length, first-strand cDNA molecules with a much higher resolution than Northern blot analysis of mRNA molecules. This method can therefore be used to discriminate between the multiple mRNA species transcribed from a given gene or the homologous mRNA species transcribed from a given gene family.

INTRODUCTION

The proportion of full-length cDNA clones in a library depends on the efficiency of synthesis of both the first- and the second-strand cDNA molecules. The most widely used method for second-strand synthesis involves digestion of the 3′ end of the first strand, thereby resulting in loss of the sequences corresponding to the 5′ end of the mRNA (9). Conversely, we previously published a method that allows one to synthesize second-strand molecules that are full-length with respect to the first-strand molecules (3). In this method, a poly(dT) tract is synthesized with terminal transferase at the 3′ end of the first strand, and the second strand is synthesized with T7 DNA polymerase using oligo(dA) as a primer. The poly(dT) tract is digested by the very high 3′ exonuclease activity of T7 DNA polymerase, except for the region annealed to the very oligo(dA) molecule used as a primer. Because the
second-strand molecules are full-length with respect to the first-strand molecules, first-strand synthesis becomes the essential limiting step for the obtention of full-length cDNA clones. In this respect, premature termination of reverse transcription (RT) is known to occur in given regions of certain mRNAs, likely because of their secondary structure. We therefore developed a method to analyze the size distribution of the first-strand molecules corresponding to the mRNA species to be cloned. As an example, we analyzed the first-strand molecules corresponding to the human c-myc mRNAs.

MATERIALS AND METHODS

Synthesis of First-Strand cDNA Molecules

Cytoplasmic polyadenylated RNA was isolated from the SW613-S human carcinoma cell line (13). RNA (1 µg) was incubated for 5 min at 70°C in 10 µL of 10 mM Tris-HCl, pH 8.0, with 0.5 µg of either phosphorylated oligo(dT)12-18 (Pharmacia Biotech, Uppsala, Sweden) or a 20-mer primer starting at nucleotide (nt) 7491 of the c-myc gene and chilled on ice. The first-strand cDNA synthesis reaction was performed for 1 h at 37°C in a volume of 50 µL containing the above mixture, 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl2, 10 mM dithiothreitol, 500 µM dATP, 500 µM dTTP, 500 µM dGTP, 500 µM dCTP, 100 µg/mL bovine serum albumin (Life Technologies, Gaithersburg, MD, USA), 1 U/µL RNasin® (Promega, Madison, WI, USA) and 200 U of either Moloney murine leukemia virus (M-MLV) Reverse Transcriptase or SUPERSCRIPT II™ Reverse Transcriptase (both from Life Technologies). After addition of EDTA to a final concentration of 20 mM, the reaction mixture was extracted with one volume of water-saturated phenol adjusted to 0.12 M Tris-HCl, pH 8.0, followed by one volume of chloroform. The organic phases were washed with one-half volume of 10 mM Tris-HCl, pH 8.0, 1 mM EDTA. After addition of sodium chloride to a final concentration of 0.1 M, the product was precipitated with two volumes of ethanol at -20°C.

Southern Blot Analysis of First-Strand cDNA Molecules

The first-strand cDNA molecules were resuspended in 10 µL of 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, and 1.5 µL of 50% glycerol, 150 mM EDTA, 0.25% bromophenol blue were added. The 1-kb DNA ladder (Life Technologies), end-labeled as described (3), was used as a molecular weight marker. A 1% agarose gel (Pharmacia Biotech) was prepared in water and incubated twice for 15 min in 2.5 vol of 30 mM NaOH, 2 mM EDTA. Electrophoresis was performed in 30 mM NaOH, 2 mM EDTA, with circulation of the buffer, at 25 V for 20 h. The gel was rinsed in water, incubated twice for 10 min in 2.5 vol of 0.25 M HCl, rinsed again in water and incubated twice for 15 min in 2.5 vol of 0.5 M NaOH, 1.5 M NaCl. Transfer was performed by capillarity in 0.5 M NaOH, 1.5 M NaCl on a GeneScreen Plus® membrane (NEN Life Science Products, Boston, MA, USA). Hybridization was performed as described (18).

Probes Specific for the Human c-myc mRNAs

The human c-myc gene is located within an 8.1-kb HindIII-EcoRI fragment, which was previously cloned from SW613-S genomic DNA (14).
The P0 and P3 promoters are located within an Apal-PvuII fragment (nt 1538–1976) and a PvuII-SmaI fragment (nt 3718–4229), respectively, which were subcloned into the pGEM®-4 plasmid (Promega). Exon 3 is almost entirely located within a Clal-EcoRI fragment (nt 6680–8078), which was previously subcloned into a plasmid (15). The recombinant plasmids were digested with HindIII and EcoRI, and the genomic fragments were purified by agarose gel electrophoresis, followed by electroelution, as described (18). The probes were prepared by random priming using a kit from Amersham International plc (Little Chalfont, Bucks, England, UK).

RESULTS

First-strand cDNA molecules were synthesized using either M-MLV Reverse Transcriptase or a mutant form devoid of RNase H activity. The mutant form is expected to synthesize full-length, first-strand molecules with a higher efficiency than the wild-type polymerase (8). Indeed, the presence of the RNase H activity probably results in competition between synthesis of the 3' end of the first-strand molecule and degradation of the 5' end of the mRNA molecule. In addition, RT was initiated using either oligo(dT) or an oligonucleotide specific for c-myc mRNAs (Figure 1). This oligonucleotide is complementary to a region of exon 3 located 19 and 160 nt upstream of the A1 and A2 polyadenylation signals, respectively. The resulting first-strand molecules were separated by electrophoresis on an alkaline agarose gel, and a Southern blot hybridization was performed using a probe encompassing most of exon 3 (probe A).

RT using oligo(dT) as a primer resulted in two bands corresponding to first-strand molecules 2.20 and 2.35 kb in length (Figure 2, lanes 2 and 4). A longer exposure resulted in a third band corresponding to first-strand molecules 2.05 kb in length (Figure 3A, lane 2). These sizes are those expected for full-length, first-strand molecules corresponding to the mRNAs transcribed from the P1 and P2 promoters, with the use of the A1 and A2 polyadenylation signals. The first-strand molecules 2.05 and 2.35 kb in length correspond to mRNAs P2A1 and P1A2, respectively (Figure 1). The first-strand molecules 2.20 kb in length correspond to mRNAs P1A1 and P2A2. RT using the oligonucleotide specific for c-myc mRNAs as a primer resulted in two bands corresponding to first-strand molecules 2.05 and 2.20 kb in length (Figure 2, lanes 3 and 5). These sizes are those expected for full-length, first-strand molecules corresponding to mRNAs transcribed from P2 and P1, respectively. Whether oligo(dT) or the oligonucleotide specific for c-myc mRNAs was used, the mutant form devoid of RNase H activity gave rise to a higher yield of full-length, first-strand molecules than the wild-type polymerase.

In addition to the exon-3 probe (probe A), which was expected to detect first-strand molecules correspond-

Figure 2. Electrophoresis on an alkaline agarose gel followed by Southern blot hybridization as a method to analyze first-strand cDNA molecules corresponding to given mRNA species. The first-strand cDNA molecules were synthesized using the M-MLV Reverse Transcriptase (lanes 2 and 3) or a mutant form devoid of RNase H activity (lanes 4 and 5). RT was initiated using either oligo(dT) (lanes 2 and 4) or an oligonucleotide specific for c-myc mRNAs (lanes 3 and 5). This oligonucleotide is complementary to a region of exon 3 located 19 and 160 nt upstream of A1 and A2 polyadenylation signals, respectively. The resulting first-strand cDNA molecules were separated by electrophoresis on an alkaline agarose gel and analyzed by Southern blot hybridization using a probe encompassing most of exon 3 (probe A). The 1-kb DNA ladder was end-labeled and used as a molecular weight marker (lane 1).

Figure 3. Southern blot analysis of first-strand cDNA molecules corresponding to human c-myc mRNAs. The first-strand cDNA molecules were synthesized using a mutant form of M-MLV Reverse Transcriptase devoid of RNase H activity and oligo(dT) as a primer. Probes encompassing most of exon 3 (probe A; Panel A), the P0 promoter (probe B; Panel B) and the P3 promoter (probe C; Panel C) were used. The 1-kb DNA ladder was end-labeled and used as a molecular weight marker (lanes 1).
ing to all c-myc mRNAs, a probe encompassing P0 (probe B) and a probe encompassing P3 (probe C) were used. Indeed, the 2.5-kb mRNA transcribed from P0 and the mRNAs transcribed from P3 were expected to co-migrate with the mRNAs transcribed from P1 and P2. Furthermore, the mRNAs transcribed from P0 and P3 were expected to be present at a much lower level than the mRNAs transcribed from P1 and P2.

Hybridization performed with the exon-3 probe (probe A) and the P0 probe (probe B) resulted in a band corresponding to first-strand molecules 2.80 to 3.10 kb in length (Figure 3, A and B, lanes 2). These sizes are those expected for full-length, first-strand molecules corresponding to the 3.1-kb mRNA transcribed from P0. Indeed, the poly(A) tracts of mRNA molecules are approximately 0.2 kb in length, P0 contains several sites of transcription initiation (2), and both A1 and A2 are probably used. Hybridization using the P0 probe (probe B) did not allow us to detect first-strand molecules corresponding to the 2.5-kb mRNA transcribed from P0. A likely explanation is that this mRNA is not transcribed in the SW613-S cell line.

Hybridization performed with the P3 probe (probe C) resulted in two bands corresponding to first-strand molecules 2.20 and 2.35 kb in length (Figure 3C, lane 2). These sizes are those expected for full-length, first-strand molecules corresponding to the mRNAs transcribed from P3, with the use of A1 and A2, respectively. The first-strand molecules 2.35 kb in length are the most abundant. Consistently, analysis of c-myc mRNAs from other human cell lines previously indicated that A2 is used more frequently than A1 (12).

First-strand molecules 3.70 to 4.05 kb in length were detected using the exon-3 probe (probe A) and the P3 probe (probe C) (Figure 3, A and C, lanes 2). In this respect, a 4.1-kb cytoplasmic RNA was previously detected using the exon-3 probe (probe A) (15). These first-strand molecules are 1.6 kb longer than those corresponding to the mRNAs transcribed from P1 and P2. In addition, the P3 probe (probe C) is located within intron 1, and intron 1 is 1624 nt in length. These first-strand molecules may therefore correspond to forms of the mRNAs transcribed from P1 and P2, which would contain intron 1. In this respect, analysis of xenopus c-myc RNAs indicated the existence of a cytoplasmic RNA containing intron 1 (11).

Finally, first-strand molecules 3.4 kb in length were detected using the exon-3 probe (probe A) (Figure 3A, lane 2). They may correspond to new mRNA species, which would remain to be characterized by cDNA cloning and
sequencing. Alternatively, they may be truncated forms of the first-strand molecules 3.7–4.05 kb in length, resulting from premature termination of RT.

**DISCUSSION**

Our method allows one to determine whether full-length, first-strand cDNA molecules corresponding to the mRNA species to be cloned are synthesized efficiently. Even if full-length molecules are synthesized efficiently, incomplete molecules are always synthesized to some extent. These molecules result from premature termination of RT or from RT of partially digested mRNA molecules. In addition, the cloning efficiency is known to be higher for shorter molecules. Size fractionation of the double-stranded cDNA molecules therefore enables one to increase the proportion of full-length cDNA clones in the library. This additional step is especially useful if the mRNA to be cloned is of low abundance or if a probe corresponding to the 5' end of the mRNA is not available. We used this approach to clone cDNAs corresponding to the human c-myc mRNAs transcribed from the four promoters (4). Most of these cDNAs (18/25) contain an additional dG residue at their 5' end, which is known to correspond to the cap of the mRNA molecules (10,19). These cDNAs are therefore full-length to the nearest nucleotide. Conversely, if full-length, first-strand molecules are not synthesized efficiently, likely because of the secondary structure of the mRNA molecules, it may be possible to circumvent the difficulty by initiating synthesis of the first-strand molecules using random primers instead of oligo(dT) molecules.

Interestingly, full-length, first-strand molecules corresponding to c-myc mRNAs differing by only 0.15 kb in length could be discriminated. Our method can therefore be used to analyze full-length, first-strand cDNA molecules with a high resolution. This is in contrast with Northern blot analysis of mRNA molecules, which usually suffers from a poor resolution. A likely explanation is that the mRNA molecules are heterogeneous in size, whereas the corresponding first-strand cDNA molecules are not. Indeed, the poly(A) tracts at the 3' ends of the mRNA molecules are known to be heterogeneous in size. Conversely, because the poly(A) tracts are saturated by oligo(dT) molecules, the first-strand cDNA molecules contain limited numbers of dT residues corresponding to the poly(A) tracts. Consistent with this view, it was previously shown that selective digestion of the poly(A) tracts of the mRNA molecules by RNase H after annealing to oligo(dT) molecules allows Northern blot analysis to be performed with a better resolution (5,20). Southern blot analysis of first-strand cDNA molecules is therefore similar to Northern blot analysis of mRNA molecules after digestion of the poly(A) tracts.

Because of its higher resolution, this method can be used to discriminate between the multiple mRNA species transcribed from a given gene, resulting from alternative splicing and alternative usage of promoters and polyadenylation signals. This method may also be used to discriminate between the homologous mRNAs transcribed from a given gene family.

**REFERENCES**

by Moloney murine leukemia virus RNAse H-
minus reverse transcriptase possessing full
DNA polymerase activity. Focus 14:91-93.
ple and very efficient method for generating
10. Hirzmann, J., D. Luo, J. Hahnen and G.
Hobom. 1993. Determination of messenger
RNA 5′-ends by reverse transcription of the
cap structure. Nucleic Acids Res. 21:3597-
3598.
alternative splicing in the Xenopus laevis
c-Myc gene creates an intron-1 containing
c-myc RNA present only in post-midblastula
Knaken, J. De Ruiter and A.J. Van Der Eb.
1989. Analysis of polyadenylation site usage
of the c-myc oncogene. Nucleic Acids Res.
17:6499-6514.
13. Lavialle, C., N. Modjtahedi, T. Lamonerie,
T. Frebourg, R.-M. Landin, N. Fossar, G.
Lhomond, R. Cassingena and O. Brison.
1989. The human breast carcinoma cell line
SW 613-S: an experimental system to study
tumor heterogeneity in relation to c-myc
amplification, growth factor production and
other markers (Review). Anticancer Res.
9:1265-1280.
1988. Cis activation of the c-myc gene in
bovine papilloma virus type 1/human c-myc
15. Modjtahedi, N., C. Lavialle, M.-F. Poupon,
R.-M. Landin, R. Cassingena, R. Monier
and O. Brison. 1985. Increased level of am-
plification of the c-myc oncogene in tumors
induced in nude mice by a human breast carci-
noma cell line. Cancer Res. 45:4372-4379.
transcript initiated in intron 1 in Friend ery-
throleukemia cells and in other murine cell
types with no evidence of c-myc gene re-
istence of a c-myc mRNA initiated in intron 1
with the normal c-myc mRNA and similar reg-
ulation of both transcripts in mammalian cells.
Oncogene Res. 5:73-78.
1989. Molecular Cloning: A Laboratory Man-
ual, 2nd ed. Cold Spring Harbor Laboratory
Press, Cold Spring Harbor, NY.
1995. Transcription of the 5′-terminal cap
nucleotide by RNA-dependent DNA poly-
merase: possible involvement in retroviral re-
verse transcription. DNA Cell Biol. 14:991-
996.
RNA processing is a limiting step for murine
tumor necrosis factor β expression in response
to interleukin-2. Mol. Cell. Biol. 10:5865-
5875.

This work was supported by grants from
the Association pour la Recherche sur le
Cancer. Address correspondence to Olivier
Brison, Laboratoire de Genetique Oncolo-
gique, CNRS URA 1967, Institut Gustave-
Roussy, 39, rue Camille-Desmoulins, 94805
Villejuif cedex, France. Internet: obrison@
gr.fr

Received 18 March 1996; accepted 2
December 1996.

Myriam Bodescot and Olivier
Brison
Institut Gustave-Roussy
Villejuif, France