Efficient Cloning of Alternatively Polyadenylated Transcripts via Hybridization Capture PCR

Theodoros N. Rampias[1], Emmanuel G.Fragoulis and Diamantis C. Sideris*

University of Athens, Faculty of Biology, Department of Biochemistry and Molecular Biology, Panepistimioupolis, 15701, Athens, Greece
[1]Present address: Biomedical Research Foundation of the Academy of Athens, Center of Basic Research II, 11527 Athens, Greece

Abstract

Cloning of alternatively polyadenylated transcripts is crucial for studying gene expression and function. Recent transcriptome analysis has mainly focused on large EST clone collections. However, EST sequencing techniques in many cases are incapable of isolating rare transcripts or address transcript variability. In most cases, 3´ RACE is applied for the experimental identification of alternatively polyadenylated transcripts. However, its application may result in nonspecific amplification and false positive products due to the usage of a single gene specific primer. Additionally, internal poly(A) stretches primed by oligo(dT) primer in mRNAs with AU-rich 3´UTR may generate truncated cDNAs. To overcome these limitations, we have developed a simple and rapid approach combining SMART technology for the construction of a full length cDNA library and hybrid capture PCR for the selection and amplification of target cDNAs. Our strategy is characterized by enhanced specificity compared to other conventional RT-PCR and 3´ RACE procedures.

Introduction

Alternative 3´ processing and polyadenylation of pre-mRNA can generate multiple transcript isoforms of the same gene and has evident roles in mRNA stability, localization and translational efficiency (Bartel, 2009; de Moor et al., 2005; Wickens et al., 2002). Despite the important role of alternative 3´ processing and polyadenylation in gene function, genome wide characterization of 3´ UTRs is still incomplete. Recent studies have shown that as many as 5,000 human genes may have unreported 3´ end extensions (Le Texier et al., 2006). Even in the well-characterized genome of *Caenorhabditis elegans*, nearly half of the genes deposited in WormBase lack an annotated 3´ UTR and only ~5% are reported to have alternative 3´ UTR isoforms (Mangone et al., 2010).

Alternative poly(A) sites may be located in the same 3´ exon and are referred to as tandem poly(A) sites or in alternative 3´exons (Yan and Marr, 2005). The presence of transcript isoforms with extended 3´ UTR is often related to mRNA stability, translational efficiency and cellular localization (Decker and Parker, 1995; Jansen, 2001; Mazumder et al., 2003; Wilusz et al., 2001). Alternative polyadenylation may

significantly alter transcript characteristics by the addition of regulatory elements within the non-coding terminal end of the transcript (Gu et al., 2009; Ji et al., 2009).

To date, the elucidation of gene expression and function requires the isolation and cloning of alternatively spliced or polyadenylated transcripts. High-throughput transcriptome research has focused on extended expressed sequence tag (EST) libraries, as well as the characterization of 3´ UTRs mainly using computational methods (Bengert and Dandekar, 2003; Chen et al., 2006). Although EST collections are a helpful tool for the identification of novel transcripts, their potential is limited by low representation of rare transcripts, poor coverage of many tissue types and developmental stages, as well as deposition of sequences from misspliced RNAs (Wang, 2008). In addition, the coverage of many organisms' transcriptome by ESTs remains incomplete due to issues such as transcript end bias, library coverage limitations and sampling differences (Modrek and Lee, 2002).

Cloning and characterization of transcript isoforms either by cDNA library screening or by rapid amplification of cDNA ends (RACE) (Frohman, 1993) is the most common method for the experimental validation of predicted extensive variations at the 3´ end of transcripts. Screening of cDNA library is an expensive, tedious and time consuming procedure. On the other hand, the conventional 3´ RACE method often has practical problems, such as low sensitivity for rare transcripts and high background of non-specific amplification products, due to internal priming of the oligo-(dT) primer in adenine rich stretches (Nam et al., 2002). Hybridization capture in solution using biotinylated DNA probes is also an alternative approach for transcript isolation from amplified (Levin et al., 2009) or enriched (Haraguchi et al., 2003) cDNA libraries, as well as for nucleic acids detection of pathogens in clinical or environmental samples (Chen et al., 1998; Jacobsen, 1995; Maibach et al., 2002). However, its application to transcriptome analysis of 3´ UTR extended isoforms remains limited.

Direct high throughput sequencing of cDNA (RNA-Seq) is a recently developed approach that uses deep-sequencing technologies, allowing the further characterization and quantification of prokaryotic and eukaryotic transcriptomes (Wang et al., 2009). RNA-Seq has also been applied to *C. elegans* transcriptome for the identification of 3´ UTR isoforms by developing a high throughput method called poly(A)-position profiling by sequencing (3P-Seq). This method is based on the isolation of the single stranded polyadenylated ends after ligation with a biotinylated adaptor at the 3´ end of mRNAs, partial digestion with RNase T_1, magnetic capture of biotinylated DNA fragments, reverse transcription - digestion with RNase H and finally, the purified polyadenylated ends are subjected to high-throughput sequencing (Jan et al., 2011). Compared to conventional methods, this approach permitted the identification of 8,580 additional UTRs in *C. elegans* transcriptome and provided evidence that thousands of deposited shorter UTR isoforms

*Corresponding author: Email: dsideris@biol.uoa.gr

that were supported by oligo(dT) based methods may not be authentic and seem to derive from internal-priming artifacts.

Although RNA-Seq is a very attractive approach to transcriptome profiling, it is still a technology under constant development and faces some technical challenges. For example, 7-nt or longer homopolymers have higher error rates, leading to ambiguous base calls in the sequence output (Huse et al., 2007; Margulies et al., 2005) and therefore the read accuracy in AU rich 3´ UTRs may be imprecise in some cases. Some technical manipulations such as RNA or cDNA fragmentation may also have a negative impact on the sequencing outcome. More specifically, RNA or cDNA molecules must be fragmented into smaller pieces to be compatible with deep-sequencing technology. RNA fragmentation has little bias over the transcript body but is relatively depleted for both the 5´ and 3´ ends. In contrary, cDNA fragmentation is strongly biased towards the identification of sequences from the 3´ end of transcripts (Wang et al., 2009). In any case, RNA-Seq technologies provide only partial sequencing information and do not allow the direct cloning of target transcripts. Therefore, EST clone collections remain the major sequence sources for many species.

Here, we present a simple and efficient approach for the experimental identification and cloning of alternatively polyadenylated transcripts with extended 3´ UTR. It combines SMART technology (Zhu et al., 2001) and hybrid capture PCR and fills an important niche in transcriptome research. We successfully applied this method for the isolation and cloning of an alternatively polyadenylated transcript of *Cc* RNase gene from the insect *Ceratitis capitata* (Rampias et al., 2008).

Materials and Methods
RNA isolation and of single stranded cDNA library construction
Total RNA was prepared from six-day-old larvae of the insect *Ceratitis capitata* by the acid guanidinium thiocyanate-phenol-chloroform method (Bouhin et al., 1992). Poly $(A)^+$ RNA was isolated from 100 µg total RNA using Dynabeads oligo $(dT)_{25}$ (Dynal AS, Oslo, Norway) in accordance with the manufacturers' protocol. Approximately 1 µg of poly $(A)^+$ RNA was reverse transcribed using the SMART PCR cDNA synthesis kit (Clonotech Laboratories, Palo Alto, CA). Briefly, a 5 µl cDNA synthesis mixture containing the poly(A)$^+$ RNA, 2.4 µM 3´SMART CDS primer IIA (5´-AAG CAG TGG TAT CAA CGC AGA GTA CT$_{(30)}$ (A/C/G/T) (A/G/C) -3´), and 2.4 µM SMART IIA oligonucleotide (5´-AAGCAGTGGTATCAACGCAGAGTACGCGGG-3´) was incubated at 72 ºC for 2 min and placed on ice for another 2 min. The reaction volume was then adjusted to 10 µl adding 2 µl 5× first-strand synthesis buffer (250 mMTris-HCl, pH 8.3, 30 mM MgCl$_2$, 375 mM KCl, 10 mM DTT), 1 µl dNTP mixture (10 mM each), 1 µl deionized water and 1 µl (200 U) of Superscript II reverse transcriptase (Life Technology, Rockville, MD). The reaction was carried out at 42 ºC for 1hour and terminated by the addition of 80 µl of TE buffer, followed by incubation at 72 ºC for 7 min. The single stranded cDNA molecules were precipitated with ethanol–sodium acetate and resuspended in 20 µl TE buffer.

Hybridization capture of target cDNAs
Hybrid selection of the *Cc* RNase target transcripts was performed using a biotin labelled single stranded cDNA probe captured on streptavidin coated magnetic beads. The capture probe was complementary to a portion of the known sequence of the Cc RNase cDNA (nucleotides 69-790, accession number AJ441124) and was synthesized by conventional PCR using the 5´ biotinylated forward primer TTGTGGAAAATCATACGAG-3´ and the reverse primer 5´-CTGCAGACATCGCTTACTT-3´. The biotinylated strand of the PCR product was attached to streptavidin coated paramagnetic beads (Dynabeads® M-280 Streptavidin, Dynal, Oslo, Norway) as recommended by the manufacturer and resuspended in 80 µl of a buffer containing 3.75 x SSC, 0.125% SDS, 1.5625 x Denhardt's.

The synthesized single stranded cDNAs (20 µl) were denatured by heating at 99°C for 5 min, followed by immediate cooling on ice for 3 min and then directly added to the biotinylated captured Cc RNase probe. The hybridization reaction (100 µl) was performed at 68 °C for 6 h in 3 x SSC, 0.1% SDS, 1.25 x Denhardt's (hybridization buffer) under continuous agitation (800 rpm) in an Eppendorf thermomixer. Following hybridization, the beads were collected by a magnetic particle separator allowing the hybridization solution to be removed and washed twice with 200 µl of 2 x SSC, 0.1% SDS, at 63°C, four times with 200 µl of 10 mM Tris-HCl (pH 8.0), 1 mM ethylanediaminetetraacetic acid (EDTA) at room temperature, and finally resuspended in 50 µl of distilled H$_2$O. All post-hybridization wash steps were performed in constant agitation of 1400 rpm using an Eppendorf thermomixer.

PCR amplification, cloning and characterization of captured cDNAs
Following hybridization and washes, the target single stranded cDNAs were amplified by PCR, using the LD primer (5´- AAG CAG TGG TAA CAA CGC AGA GT-3´), which is complementary to the SMART adaptor at the 3´ and 5´ end of the synthesized cDNAs. Ten microliters of purified hybrids and 5 U of *Pfu* Turbo DNA polymerase (Stratagene, CA, USA) were used for the PCR reaction, performed at 95°C for 3 min, followed by 25 cycles of 95°C for 1 min, 56°C for 1 min, 68°C for 2 min and a final extension of 10 min at 68°C.

The amplified products were then purified using the QIAquick PCR purification kit (Qiagen), cloned into the pCR2.1 vector (Invitrogen, Carlsbad, CA, USA) and characterized by restriction endonuclease mapping and DNA sequencing.

3´ RACE-PCR
One µg poly $(A)^+$ RNA was reverse transcribed using the oligo(dT) adaptor primer (5´-GGCCACGCGTCGACTAGTACT$_{18}$-3´) and 1 µl of of the single stranded cDNA library was used as template in a PCR reaction, containing the gene specific primer DS74 (5´-TTGTGGAAAATCATACGAGA-3´) and the adaptor primer (5´-GGCCACGCGTCGACTAGTAC-3´). Ten microliters of the PCR products were analyzed on an ethidium bromide-stained 1% agarose gel and subsequently employed for Southern blot analysis.
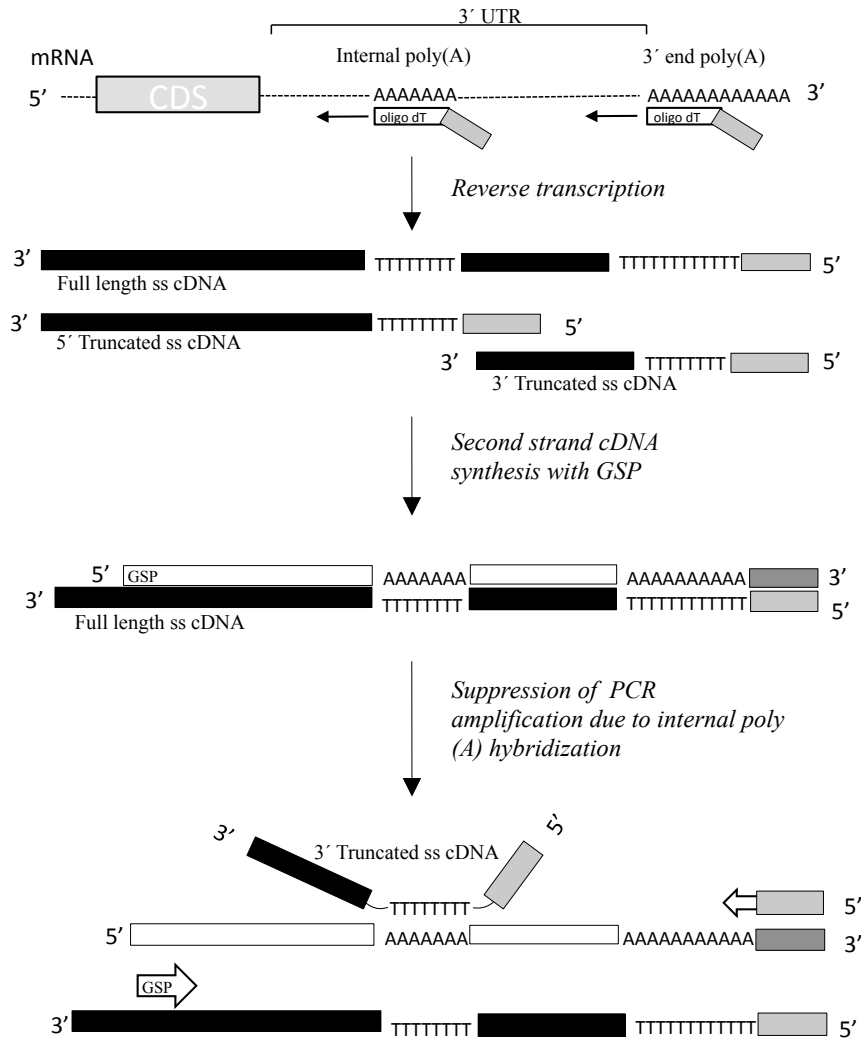
Figure 1. Schematic representation of 3′ RACE PCR suppression in mRNAs containing internal poly(A) stretches. The initiation of cDNA synthesis from internal poly(A) stretches located in 3′ UTR results in the generation of 5′ and 3′ truncated single stranded cDNAs from a single mRNA template. A gene specific primer (GSP) is used for the second strand cDNA synthesis and the double stranded cDNA molecules are then amplified in a PCR reaction using the gene specific primer and the adaptor one. During amplification, the population of 3′ truncated single stranded cDNA molecules generated in RT reaction prevents the extension of the adaptor primer due to hybridization of internal poly(A) stretches and therefore suppresses the full length cDNA synthesis.

Southern blot analysis
Southern blot experiments were carried out as described previously (Maniatis et al., 1989), using a ^{32}P-labeled DNA probe corresponding to nucleotides 58-784 of the Cc cDNA sequence (Accession number AJ441124).

Results and Discussion
The extended UTR region in alternatively polyadenylated transcripts is almost 90% longer than the constitutive UTR region (cUTR) and has a higher AU content (Tian et al., 2005). These AU rich sequences often contain poly (A) stretches which can be primed by oligo(dT) adaptor-primer during the reverse transcription reaction, resulting in the generation of 3′- and 5′- truncated single stranded cDNA molecules (Nam et al., 2002) and in the accumulation of non-specific 3′ RACE amplification products. Moreover, since the 3′- truncated single stranded cDNAs contain oligo(dT) sequences, they may be hybridized to internal poly(A) stretches of the cDNA template, suppressing the full length target cDNA synthesis (Figure 1).

To overcome this problem, we have developed a simple approach in order to increase the specificity and the efficiency of the target cDNA amplification.

The basic principle of the method is outlined in figure 2. It consists of the following steps: (i) Construction of a single stranded cDNA library according to the SMART protocol. (ii) Hybridization of cDNA molecules with a biotinylated single strand DNA probe immobilized on streptavidin paramagnetic beads. (iii) Magnetic pull down and wash of the immobilized hybrids. (iv) PCR amplification of the target
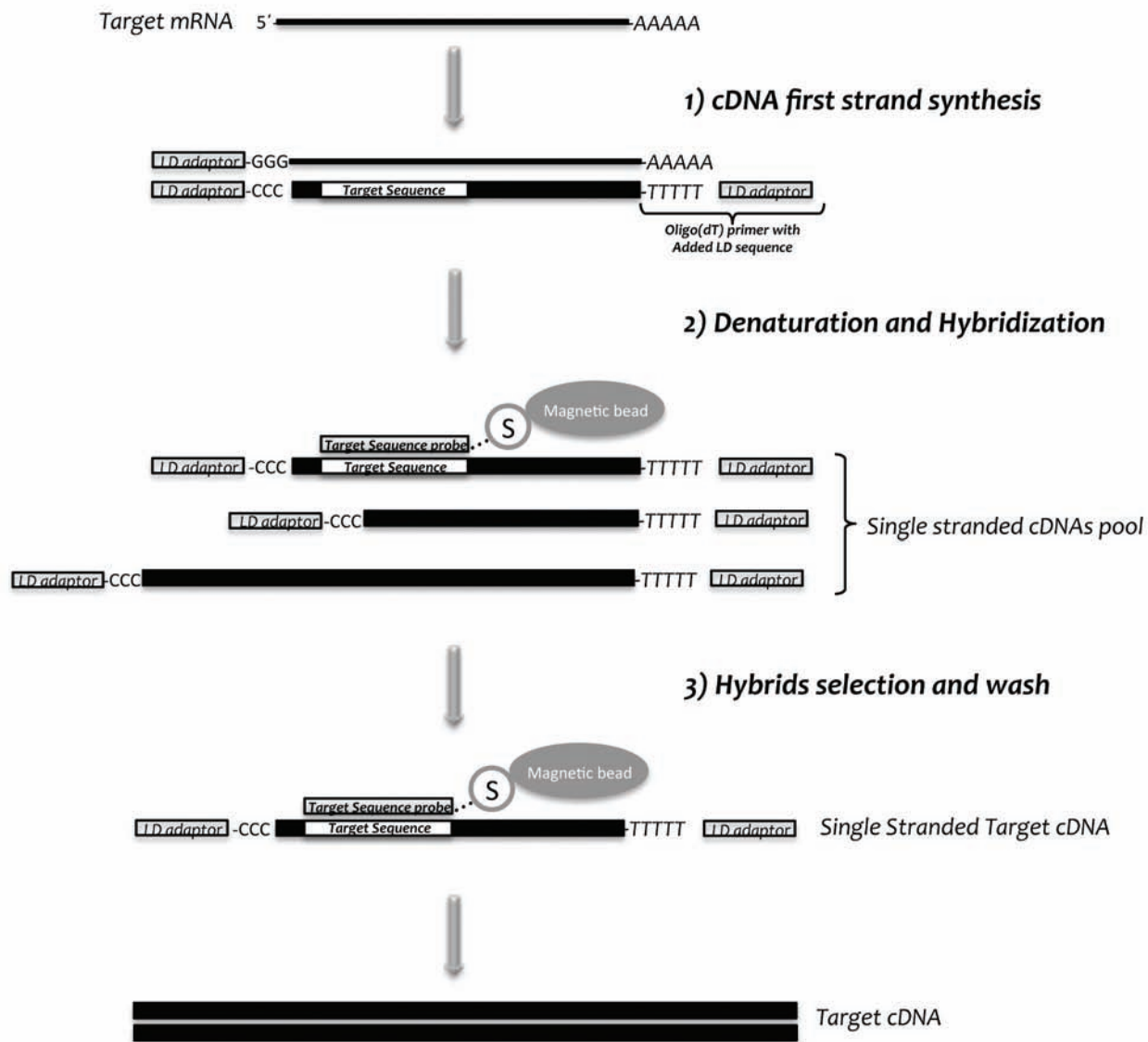
Figure 2. Schematic representation of the hybrid capture protocol for the selection of alternative transcripts. mRNAs are represented by black line and cDNAs by black boxes.

cDNA molecules, using a primer complementary to the 5′ SMART adaptor sequence (LD primer).

To our knowledge this is the first adapted magnetic capture hybridization PCR protocol to study the expression of alternative transcripts by employing the SMART cDNA synthesis technology.

Construction of cDNA library and hybrid selection of target transcripts
In order to evaluate the efficiency of our protocol, we prepared a single stranded SMART cDNA library from 6-day-old larvae of the insect *Ceratitis capitata*. SMART technology offers the advantage of full length cDNA synthesis, due to the usage of the SMART oligo that specifically binds at the 5′ end of mRNA molecules. The SMART anchor sequence and the poly(A) sequence serve as universal priming sites for end-to-end cDNA amplification. This approach eliminates the amplification of 5′or 3′ truncated molecules.

In the next step, the isolation of the Cc RNase target transcripts was performed by hybrid selection using a biotin-labelled single stranded DNA probe captured on streptavidin coated magnetic beads. The DNA probe is designed to be complementary to a part of the coding region of the Cc RNase gene and has a length of 721 nucleotides and a melting temperature (Tm) of 60-70°C. The synthesized single stranded SMART cDNA library was hybridized at 68°C for 6 hours under continuous agitation to ensure that magnetic beads remained in suspension. The enrichment of target transcripts is especially crucial when working with low copy alternative transcripts. High stringency hybridization conditions facilitates the removal of the majority of non target cDNA molecules, increasing the concentration of the specific target transcripts by several orders of magnitude. Previous reports suggest that hybridization temperature should not exceed 70°C, so as to reduce the probability of bead degradation (St John and Quinn, 2008). In our case

the Dynal beads tolerated a hybridization temperature of 68°C making them suitable for this protocol.

Isolation and characterization of alternative isoforms
Following hybridization, the beads were collected by a magnetic particle separator allowing the hybridization solution to be removed and washed twice with 2 x SSC, 0.1% SDS, at 63°C and four times with TE buffer at room temperature. The selected ss cDNAs captured on the beads were used as a template in polymerase chain reaction with the LD primer, that specifically binds at both ends of the cDNA molecules. The application of the above strategy in our case, allowed the retrieval of two distinct bands, as visualized in a 1.2% agarose gel (Figure 3A, lane 1). These cDNAs correspond to the two transcripts of the *Cc* RNase gene (Rampias et al., 2008), as confirmed by Southern blot (Figure 3B, lane 1) and sequencing analysis. These results demonstrated the high specificity of transcript selection and amplification in our method.

A series of experiments under different salt concentrations (2 to 3.5 × SSC in hybridization buffer) was performed in order to determine the optimal hybridization conditions. Comparison of the final PCR amplification products revealed that the usage of 3 x SSC hybridization buffer assured the retrieval of the most abundant and specific DNA product (Figure 4A).

The capture of hybrids on paramagnetic beads also allows extensive washing prior to the final PCR reaction. Applying different washing conditions of the selected beads, demonstrated that the stringency of wash is a crucial parameter, since even traces of non-target DNA can produce severe background in the subsequent PCR amplification step. When beads were initially washed with 2 × SSC containing 0.1% SDS at hybridization temperature (63°C) and then with TE at room temperature, an optimal elimination of non-specific cDNA molecules at the final amplification step was achieved. In contrary, when higher salt concentration in wash buffer (3-3.5 x SSC) or lower incubation temperatures (55-60°C) were applied, the non specific amplification was significantly increased (Figure 4B).

A relatively low number of amplification cycles (25 cycles) were applied in order to avoid PCR bias. However, a higher number of cycles can be used under optimized conditions to further increase the sensitivity of the proposed method. Moreover, the hybridization beads may also be used as a direct template for real time PCR amplification using specific primers for each transcript isoform. This approach could provide a direct and accurate measurement of the expression levels of the transcript variants.

The development of this technique was attempted due to the difficulty of isolating the alternative *Cc* RNase transcript applying the conventional 3′ RACE procedure (Figure 3A, lane 2). As shown in Figure 3, only specific PCR products were detected following our method, whereas the application of the conventional 3′ RACE (Frohman, 1993) led to the accumulation of non-specific products and failed to specifically amplify the longer transcript. Sequence
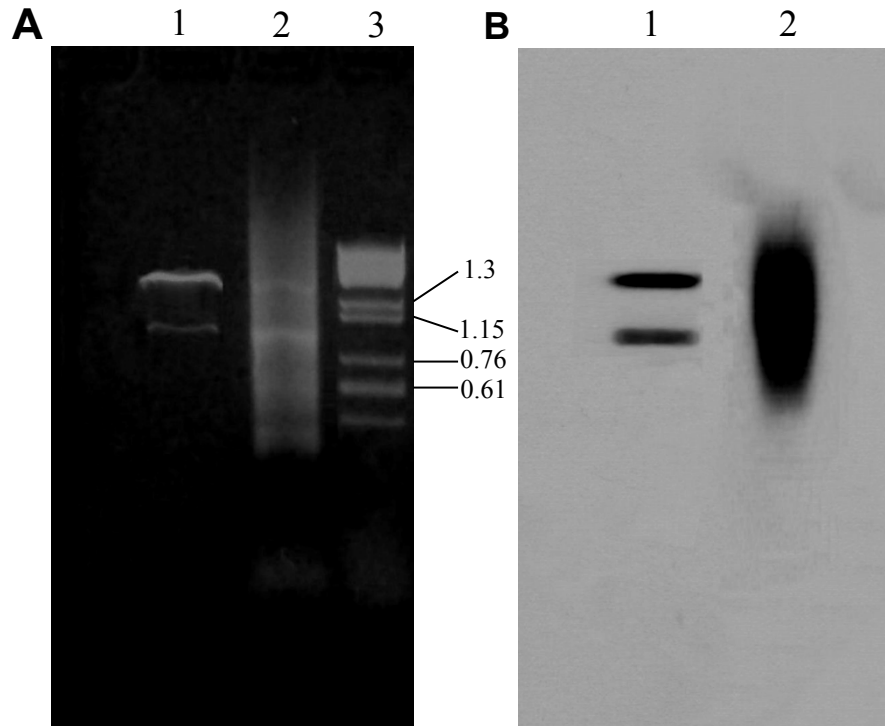

Figure 3. Comparative demonstration of the PCR products obtained applying the magnetic capture hybrid selection technique (lane 1) and 3′ RACE (lane 2). (A) The PCR products were analyzed on an ethidium bromide-stained 1.2% agarose gel in parallel with λ phage DNA digested with *Hin*dIII (lane 3). (B) Southern blot analysis of PCR amplification products with a ^{32}P-labelled specific probe corresponding to nucleotides 69-790 of the *Cc* RNase gene.
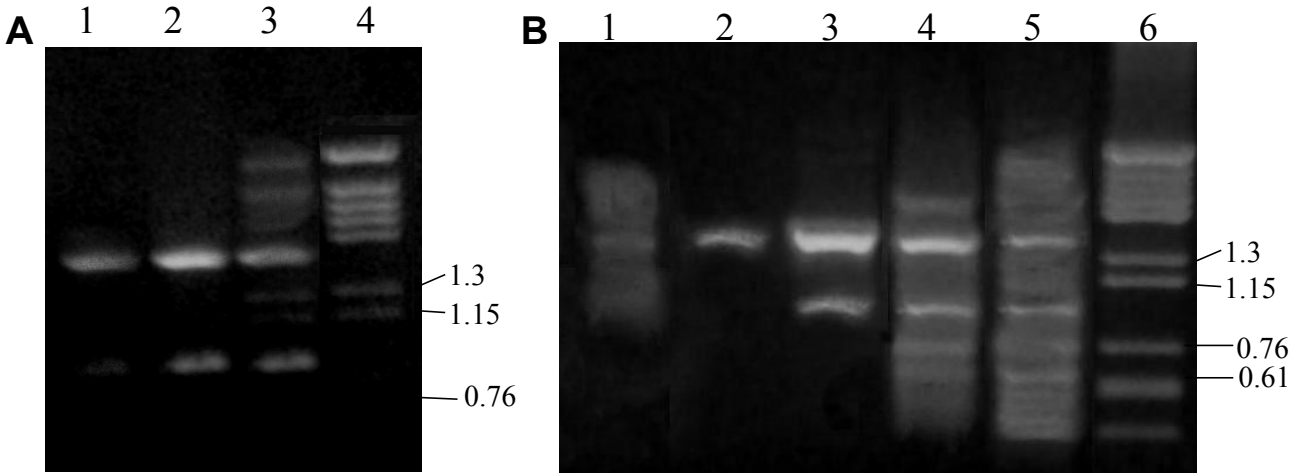
A

1.3
1.15
0.76

B

1.3
1.15
0.76
0.61

Figure 4. Comparative demonstration of the PCR products obtained under several hybridization and washing conditions. (A) The final PCR products yielded under different hybridization conditions, 2 x SSC (lane 1), 3 x SSC (lane 2) and 3.5 x SSC (lane 3) and (B) under different washing conditions 2 x SSC at 55 °C (lane 1), 2 x SSC at 60 °C (lane 2), 2 x SSC at 63 °C (lane 3), 3 x SSC at 63 °C (lane 4) and 3.5 x SSC at 63 °C (lane 5) were analyzed on an ethidium bromide-stained 1.2% agarose gel in parallel with λ phage DNA digested with *Hind*III (lane 4A and lane 6B).

analysis of the longer Cc RNase transcript revealed a 71% AU content, as well as the presence of several adenine-rich stretches in the extended UTR region (Ac. No AJ 874689). It is possible that the internal priming of the oligo-(dT) containing primer to these adenine-rich stretches during standard 3′ RACE led to non specific amplification. This hypothesis is in agreement with previous reports demonstrating that the high rate of oligo-(dT) priming in internal poly(A) tracts during reverse transcription can generate a high background of nonspecific products (Nam et al., 2002). In our protocol, the addition of the hybridization capture step allows the complete removal of the 3′-truncated ss cDNA molecules prior to the final amplification step, facilitating the complete target cDNA synthesis and allowing the identification of transcript variants with adenine rich regions in their sequence. Given the fact that the high AU content of the extended 3′ untranslated region is a general feature of the alternatively polyadenylated transcripts, our method is suitable for the efficient amplification of transcript variants generated by alternative choice of the polyadenylation signal.
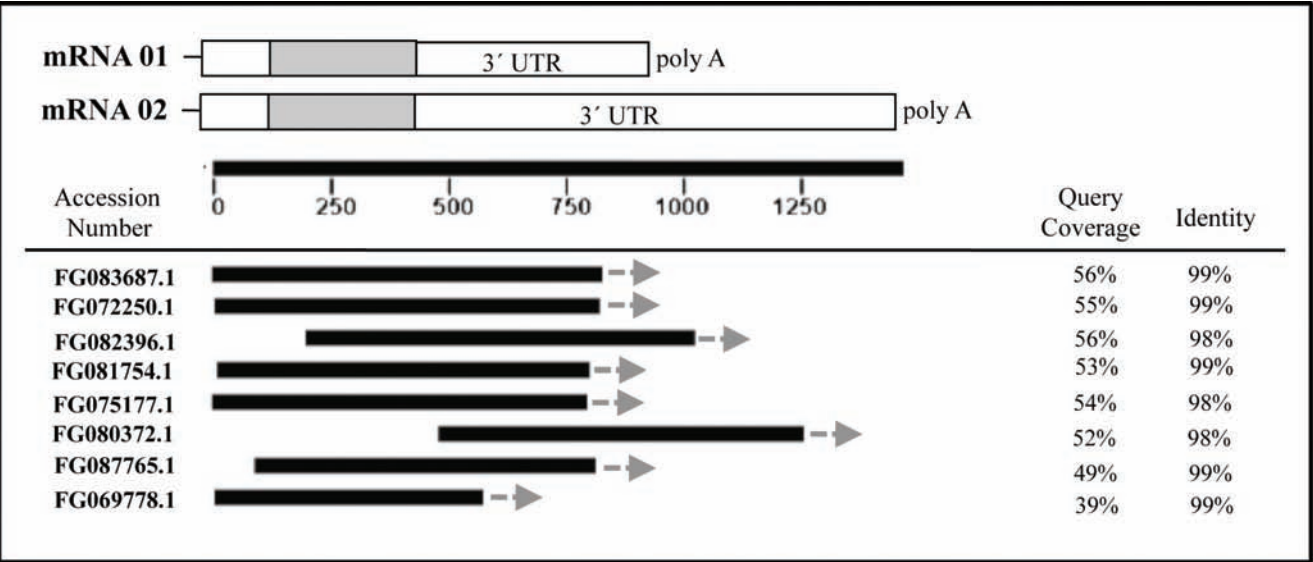

Figure 5. BLASTN analysis of the Cc RNase transcripts (mRNA 01, mRNA 02) using the *Ceratitis capitata* NCBI dbEST. Each EST is represented by a horizontal line. Arrows show the direction of EST sequencing read. The accession numbers of corresponding ESTs, the identity and the coverage of the query sequence are also indicated. The accession numbers of the Cc RNase transcripts are AJ441124 (mRNA 01) and AJ874689 (mRNA 02).

Since our previously reported experimental results demonstrated that the longer Cc RNase transcript is more abundant than the shorter one (Rampias et al., 2008), a question about the representation of Cc RNase transcripts in the current expression databases is raised. Recently, the first major EST dataset from the insect *Ceratitis capitata* embryo and adult head cDNA libraries was released (Gomulski et al., 2008). A BLASTN analysis using the longer Cc RNase transcript sequence as query against this database retrieved six ESTs corresponding to the full length sequence of the shorter Cc RNase transcript and only two ESTs (FG082396.1 and FG080372.1) corresponding to 3´- truncated sequences of the longer transcript (Figure 5). In addition, given that all the retrieved ESTs represented 5´ sequencing reads of the cDNA clones, the BLASTN analysis pointed out the lack of currently available sequence information regarding the 3´ end of the longer Cc RNase transcript in EST databases. In fact, the method described in this report could avert this incompleteness since it facilitates the isolation of full-length 3´ UTR extended transcripts that are not represented in EST databases.

The protocol presented here is highly flexible and can also have potential applications in the identification of transcripts that contain regulatory elements within their 3´ UTR, such as AU-rich elements (AREs) (Barreau et al., 2005) or retroelement insertions (Lee et al., 2008), by designing DNA hybridization probes highly specific for these elements. It is well known that some 3´ UTR regulatory elements are extremely conserved reflecting a very strong selective pressure for these sequences among different transcripts. Alternative polyadenylation produces transcript isoforms with 3´ untranslated regions of different lengths. Therefore, if a regulatory element is present in an extended UTR, only those element-containing transcript isoforms are regulated (Legendre et al., 2006). Recently, it was shown that proliferating cells such as CD4[+] lymphocytes express mRNAs with shortened 3´ untranslated regions with fewer microRNA target sites (Sandberg et al., 2008). Thereafter, we believe that our protocol can also be adapted to detect alternative transcripts that host regulatory UTR elements and microRNA target sites by using the appropriate cDNA probes.

Acknowledgements

References

Barreau, C., Paillard, L., and Osborne, H.B. (2005). AU-rich elements and associated factors: are there unifying principles? Nucleic Acids Res 33, 7138-7150.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell 136, 215-233.

Bengert, P., and Dandekar, T. (2003). A software tool-box for analysis of regulatory RNA elements. Nucleic Acids Res 31, 3441-3445.

Bouhin, H., Charles, J.P., Quennedey, B., and Delachambre, J. (1992). Developmental profiles of epidermal mRNAs during the pupal-adult molt of Tenebrio molitor and isolation of a cDNA clone encoding an adult cuticular protein: effects of a juvenile hormone analogue. Dev Biol 149, 112-122.

Chen, J., Johnson, R., and Griffiths, M. (1998). Detection of verotoxigenic Escherichia coli by magnetic capture-hybridization PCR. Appl Environ Microbiol 64, 147-152.

Chen, Z., Wang, W., Ling, X.B., Liu, J.J., and Chen, L. (2006). GO-Diff: mining functional differentiation between EST-based transcriptomes. BMC Bioinformatics 7, 72.

de Moor, C.H., Meijer, H., and Lissenden, S. (2005). Mechanisms of translational control by the 3' UTR in development and differentiation. Semin Cell Dev Biol 16, 49-58.

Decker, C.J., and Parker, R. (1995). Diversity of cytoplasmic functions for the 3' untranslated region of eukaryotic transcripts. Curr Opin Cell Biol 7, 386-392.

Frohman, M.A. (1993). Rapid amplification of complementary DNA ends for generation of full-length complementary DNAs: thermal RACE. Methods Enzymol 218, 340-356.

Gomulski, L.M., Dimopoulos, G., Xi, Z., Soares, M.B., Bonaldo, M.F., Malacrida, A.R., and Gasperi, G. (2008). Gene discovery in an invasive tephritid model pest species, the Mediterranean fruit fly, Ceratitis capitata. BMC Genomics 9, 243.

Gu, S., Jin, L., Zhang, F., Sarnow, P., and Kay, M.A. (2009). Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs. Nat Struct Mol Biol 16, 144-150.

Haraguchi, S., Tsuda, M., Kitajima, S., Sasaoka, Y., Nomura-Kitabayashid, A., Kurokawa, K., and Saga, Y. (2003). nanos1: a mouse nanos gene expressed in the central nervous system is dispensable for normal development. Mech Dev 120, 721-731.

Huse, S.M., Huber, J.A., Morrison, H.G., Sogin, M.L., and Welch, D.M. (2007). Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biol 8, R143.

Jacobsen, C.S. (1995). Microscale detection of specific bacterial DNA in soil with a magnetic capture-hybridization and PCR amplification assay. Appl Environ Microbiol 61, 3347-3352.

Jan, C.H., Friedman, R.C., Ruby, J.G., and Bartel, D.P. (2011). Formation, regulation and evolution of Caenorhabditis elegans 3'UTRs. Nature 469, 97-101.

Jansen, R.P. (2001). mRNA localization: message on the move. Nat Rev Mol Cell Biol 2, 247-256.

Ji, Z., Lee, J.Y., Pan, Z., Jiang, B., and Tian, B. (2009). Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc Natl Acad Sci U S A 106, 7028-7033.

Le Texier, V., Riethoven, J.J., Kumanduri, V., Gopalakrishnan, C., Lopez, F., Gautheret, D., and Thanaraj, T.A. (2006). AltTrans: transcript pattern variants annotated for both alternative splicing and alternative polyadenylation. BMC Bioinformatics 7, 169.

Lee, J.Y., Ji, Z., and Tian, B. (2008). Phylogenetic analysis of mRNA polyadenylation sites reveals a role of transposable elements in evolution of the 3'-end of genes. Nucleic Acids Res 36, 5581-5590.

Legendre, M., Ritchie, W., Lopez, F., and Gautheret, D. (2006). Differential repression of alternative transcripts: a screen for miRNA targets. PLoS Comput Biol 2, e43.

Levin, J.Z., Berger, M.F., Adiconis, X., Rogov, P., Melnikov, A., Fennell, T., Nusbaum, C., Garraway, L.A., and Gnirke, A. (2009). Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts. Genome Biol 10, R115.

Maibach, R.C., Dutly, F., and Altwegg, M. (2002). Detection of Tropheryma whipplei DNA in feces by PCR using a target capture method. J Clin Microbiol 40, 2466-2471.

Mangone, M., Manoharan, A.P., Thierry-Mieg, D., Thierry-Mieg, J., Han, T., Mackowiak, S.D., Mis, E., Zegar, C., Gutwein, M.R., Khivansara, V., et al. (2010). The landscape of C. elegans 3'UTRs. Science 329, 432-435.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1989). Molecular cloning : a laboratory manual, 2nd edn (Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory).

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., et al. (2005). Genome sequencing in open microfabricated high-density picolitre reactors. Nature 437, 376-380.

Mazumder, B., Seshadri, V., and Fox, P.L. (2003). Translational control by the 3'-UTR: the ends specify the means. Trends Biochem Sci 28, 91-98.

Modrek, B., and Lee, C. (2002). A genomic view of alternative splicing. Nat Genet 30, 13-19.

Nam, D.K., Lee, S., Zhou, G., Cao, X., Wang, C., Clark, T., Chen, J., Rowley, J.D., and Wang, S.M. (2002). Oligo(dT) primer generates a high frequency of truncated cDNAs through internal poly(A) priming during reverse transcription. Proc Natl Acad Sci U S A 99, 6152-6156.

Rampias, T.N., Fragoulis, E.G., and Sideris, D.C. (2008). Genomic structure and expression analysis of the RNase kappa family ortholog gene in the insect Ceratitis capitata. Febs J 275, 6217-6227.

Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., and Burge, C.B. (2008). Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science 320, 1643-1647.

St John, J., and Quinn, T.W. (2008). Rapid capture of DNA targets. Biotechniques 44, 259-264.

Tian, B., Hu, J., Zhang, H., and Lutz, C.S. (2005). A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res 33, 201-212.

Wang, S.M. (2008). Long-short-long games in mRNA identification: the length matters. Curr Pharm Biotechnol 9, 362-367.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10, 57-63.

Wickens, M., Bernstein, D.S., Kimble, J., and Parker, R. (2002). A PUF family portrait: 3'UTR regulation as a way of life. Trends Genet 18, 150-157.

Wilusz, C.J., Wormington, M., and Peltz, S.W. (2001). The cap-to-tail guide to mRNA turnover. Nat Rev Mol Cell Biol 2, 237-246.

Yan, J., and Marr, T.G. (2005). Computational analysis of 3'-ends of ESTs shows four classes of alternative polyadenylation in human, mouse, and rat. Genome Res 15, 369-375.

Zhu, Y.Y., Machleder, E.M., Chenchik, A., Li, R., and Siebert, P.D. (2001). Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. Biotechniques 30, 892-897.
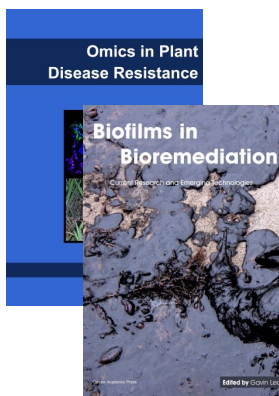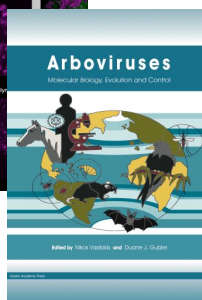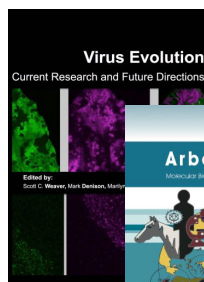
Further Reading

- **MALDI-TOF Mass Spectrometry in Microbiology**
  Edited by: M Kostrzewa, S Schubert (2016)
  www.caister.com/malditof

- *Aspergillus* and *Penicillium* in the Post-genomic Era
  Edited by: RP Vries, IB Gelber, MR Andersen (2016)
  www.caister.com/aspergillus2

- **The Bacteriocins: Current Knowledge and Future Prospects**
  Edited by: RL Dorit, SM Roy, MA Riley (2016)
  www.caister.com/bacteriocins

- **Omics in Plant Disease Resistance**
  Edited by: V Bhadauria (2016)
  www.caister.com/opdr

- **Acidophiles: Life in Extremely Acidic Environments**
  Edited by: R Quatrini, DB Johnson (2016)
  www.caister.com/acidophiles

- **Climate Change and Microbial Ecology: Current Research and Future Trends**
  Edited by: J Marxsen (2016)
  www.caister.com/climate

- **Biofilms in Bioremediation: Current Research and Emerging Technologies**
  Edited by: G Lear (2016)
  www.caister.com/biorem

- **Microalgae: Current Research and Applications**
  Edited by: MN Tsaloglou (2016)
  www.caister.com/microalgae

- **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**
  Edited by: H Shintani, A Sakudo (2016)
  www.caister.com/gasplasma

- **Virus Evolution: Current Research and Future Directions**
  Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
  www.caister.com/virusevol

- **Arboviruses: Molecular Biology, Evolution and Control**
  Edited by: N Vasilakis, DJ Gubler (2016)
  www.caister.com/arbo

- *Shigella*: **Molecular and Cellular Biology**
  Edited by: WD Picking, WL Picking (2016)
  www.caister.com/shigella

- **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**
  Edited by: AM Romaní, H Guasch, MD Balaguer (2016)
  www.caister.com/aquaticbiofilms

- **Alphaviruses: Current Biology**
  Edited by: S Mahalingam, L Herrero, B Herring (2016)
  www.caister.com/alpha

- **Thermophilic Microorganisms**
  Edited by: F Li (2015)
  www.caister.com/thermophile
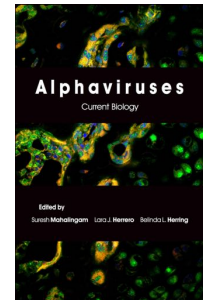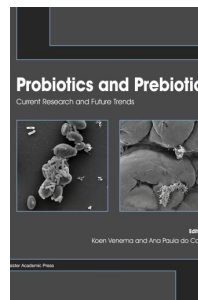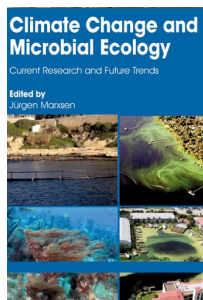
- **Flow Cytometry in Microbiology: Technology and Applications**
  Edited by: MG Wilkinson (2015)
  www.caister.com/flow

- **Probiotics and Prebiotics: Current Research and Future Trends**
  Edited by: K Venema, AP Carmo (2015)
  www.caister.com/probiotics

- **Epigenetics: Current Research and Emerging Trends**
  Edited by: BP Chadwick (2015)
  www.caister.com/epigenetics2015

- *Corynebacterium glutamicum*: **From Systems Biology to Biotechnological Applications**
  Edited by: A Burkovski (2015)
  www.caister.com/cory2

- **Advanced Vaccine Research Methods for the Decade of Vaccines**
  Edited by: F Bagnoli, R Rappuoli (2015)
  www.caister.com/vaccines

- **Antifungals: From Genomics to Resistance and the Development of Novel Agents**
  Edited by: AT Coste, P Vandeputte (2015)
  www.caister.com/antifungals

- **Bacteria-Plant Interactions: Advanced Research and Future Trends**
  Edited by: J Murillo, BA Vinatzer, RW Jackson, et al. (2015)
  www.caister.com/bacteria-plant

- *Aeromonas*
  Edited by: J Graf (2015)
  www.caister.com/aeromonas

- **Antibiotics: Current Innovations and Future Trends**
  Edited by: S Sánchez, AL Demain (2015)
  www.caister.com/antibiotics

- *Leishmania*: **Current Biology and Control**
  Edited by: S Adak, R Datta (2015)
  www.caister.com/leish2

- *Acanthamoeba*: **Biology and Pathogenesis (2nd edition)**
  Author: NA Khan (2015)
  www.caister.com/acanthamoeba2

- **Microarrays: Current Technology, Innovations and Applications**
  Edited by: Z He (2014)
  www.caister.com/microarrays2

- **Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications**
  Edited by: D Marco (2014)
  www.caister.com/n2