Analytical Biochemistry 383 (2008) 307–310

Contents lists available at ScienceDirect

Analytical Biochemistry

journal homepage: www.elsevier.com/locate/yabio

Identification of low-abundance alternatively spliced mRNA variants by exon exclusive reverse transcriptase polymerase chain reaction

Feng Wang [a], Yong Zhao [a,1], Yuhua Hao [a], Zheng Tan [a,b,*]

[a] State Key Laboratory of Biomembrane and Membrane Biotechnology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, People's Republic of China
[b] Laboratory of Biochemistry and Biophysics, College of Life Sciences, Wuhan University, Wuhan 430072, People's Republic of China

ARTICLE INFO

Article history:
Received 23 July 2008
Available online 10 September 2008

Keywords:
Alternative splicing
Exons
mRNA
Reverse transcriptase polymerase chain reaction
Abundance

ABSTRACT

Alternative splicing of messenger RNA (mRNA) precursors generates multiple transcripts from a single primary transcript. Identification and verification of splice variants and cloning of the corresponding isoforms is crucial for analyzing gene expression and understanding the related functions. For a specific gene, the abundance of the transcripts produced can vary significantly and is subject to various regulations. It can be difficult to detect low-level splicing variants when others are present in high abundance. Here we describe a method for the amplification of low-abundance mRNA splicing variants for such situations. This method introduces a hydrolysis step prior to the conventional reverse transcriptase polymerase chain reaction (RT–PCR). After the transcripts are reverse-transcripted into complementary DNA (cDNA), the cDNA of high-abundance transcript is suppressed from amplification by cleavage at the chosen exon to enhance the amplification of the low-abundance transcripts that do not have the targeted exon and are normally undetectable. We provide two examples to illustrate the detection of low-abundance splicing variants from two genes.

Alternative splicing is a posttranscriptional mechanism that makes a crucial contribution to the diversity of proteins. Genome-wide analyses indicate that 40 to 60% [1] or more [2] of human genes have alternative spliced forms. On average, a human gene contains 8.8 exons separated by 7.8 introns [3]. By selective inclusion and exclusion of different exons, alternative splicing generates multiple transcripts from a single coding sequence. Identification, verification, and cloning of messenger RNA (mRNA)[2] transcripts are important for understanding the function of the corresponding gene.

It is common that alternative splicing generates dozens of different transcripts (or even more). The abundance of transcripts can vary significantly [4] and is subject to various spatial/temporal regulations in coordination with diverse cellular functions in both physiological and pathological situations [5]. Reverse transcriptase polymerase chain reaction (RT–PCR) followed by sequencing and cloning plays a significant role in identifying and verifying novel splicing isoforms [6]. However, it is difficult to detect relatively low-level splicing isoforms when others are present in high abundance and amplified simultaneously.

In this work, we describe a method for the amplification of low-abundance RNA splicing variants. The method introduces a hydrolysis step to the conventional RT–PCR protocol to destroy the complementary DNA (cDNA) of high-abundance RNA splicing variants (Fig. 1). After transcripts are reverse-transcripted into cDNA, the cDNA of high-abundance transcript is suppressed from amplification by hydrolysis at the chosen exon. This treatment significantly reduces the competition of the high-abundance transcript and allows the low-abundance transcripts that do not have the chosen exon to be amplified. We provide examples to show how low-abundance transcripts that originally could not be detected were efficiently detected and identified.

Materials and methods

RNA isolation and reverse transcription

Total RNA was purified from cultured cells as described previously [7]. Purified RNA (1 μg) was treated with 2 U DNase I for 15 min at 37 °C in 40 mM Tris–HCl (pH 7.5), 8 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 50 U Rasin, and 1 mM Mn^{2+} in a volume of 10 μl [8]. For the first-strand cDNA synthesis, the RNA was mixed with 100 ng poly(T) primers on the ice, and the mixture was incubated at 70 °C for 5 min and then quickly chilled on the ice to avoid high-order structure formation. Reverse transcription was carried

* Corresponding author. Fax: +86 10 6480 7099.
E-mail addresses: z.tan@ioz.ac.cn, tanclswu@public.wh.hb.cn (Z. Tan).

[1] Present address: Department of Cell Biology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.

[2] *Abbreviations used:* mRNA, messenger RNA; RT–PCR, reverse transcriptase polymerase chain reaction; cDNA, complementary DNA; DTT, dithiothreitol; dNTP, deoxynucleoside triphosphate; MMLV, Moloney murine leukemia virus; BSA, bovine serum albumin; ORF, open reading frame; GPI, glycosylphosphatidylinositol; ASAP, Alternative Splicing Annotation Project.

0003-2697/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2008.09.002

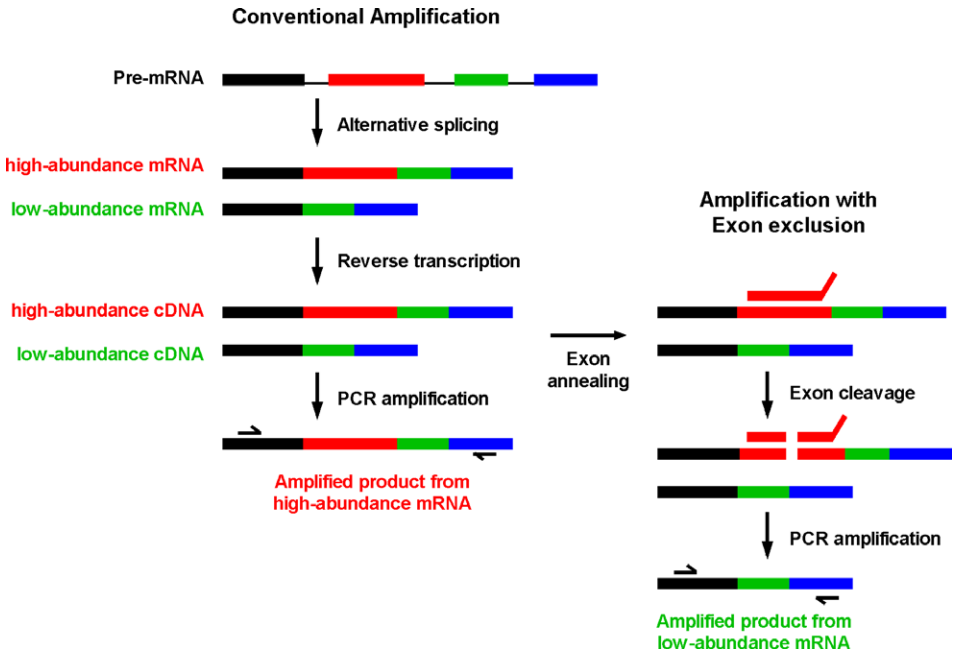

Fig. 1. Scheme of the exon exclusive RT–PCR. After reverse transcription, a deoxyoligonucleotide is annealed to a selected exon in the cDNA of the high-abundance transcript, followed by cleavage with a proper endonuclease. The cleavage of the exon prevents or suppresses the amplification of the high-abundance transcript and favors that of those low-abundance transcripts that do not carry the cleaved exon. The two primers are placed at the 5′ and 3′ sides of the cleaved exon (not necessarily the first and last exons as shown), respectively. Two unpaired nucleotides are introduced to the 3′ end of the exon cleavage oligonucleotide to prevent it from participating in PCR.

out at 42 °C for 1 h in 1× RT buffer (50 mM Tris–HCl [pH 8.3], 75 mM KCl, 3 mM $MgCl_2$, and 10 mM DTT) containing 2 mM deoxynucleoside triphosphate (dNTP), 100 U Moloney murine leukemia virus (MMLV) reverse transcriptase, and 50 U Rasin and was terminated by heating at 90 °C for 5 min according to the manufacturer's instructions.

Endonuclease digestion

First-strand cDNA from 1 μg total RNA was mixed with synthetic deoxyoligonucleotide (Table 1) complementary to the intended exon with two 3′ unpaired nucleotides to a final concentration of 1 μM in a volume of 20 μl. The mixture was heated at 94 °C for 2 min, 85 °C for 15 min, 70 °C for 15 min, 55 °C for 15 min, and room temperature for 15 min. Then 10 to 15 U restriction endonuclease was added, and samples were incubated at 37 °C for the indicated periods in digestion buffer supplied by the manufacturers (50 mM Tris–HCl [pH 7.5], 10 mM $MgCl_2$, 1 mM DTT, and 100 mM NaCl for *Xho*I; 33 mM Tris–acetate [pH 7.9], 10 mM Mg acetate, 0.5 mM DTT, 66 mM K acetate, and 0.01% bovine serum albumin [BSA] for *Bsp*1407I and *Msp*I; 10 mM Tris–HCl [pH 7.5], 10 mM $MgCl_2$, and 1 mM DTT for *Sac*I).

Table 1
Deoxyoligonucleotides used in exon cleavage and PCR

Target		Sequence
Fas	Upstream primer	5′-ATGCTGGGCATCTGGACCCT-3′
	Downstream primer	5′-AAAATCTACAAATATGTTGGCTCTT-3′
	Exon 4 cleavage (*Msp*I)	5′-GGAAATAAACTGCACCCGGACCCA GAATACCAcc-3′
PIG-T	Upstream primer	5′-TATGCCGCTTGCTCTGCTCG-3′
	Downstream primer	5′-TGGCGTGTACTCGGTCCACTTC-3′
	Exon 3 cleavage (*Sac*I)	5′-ATCTTGGAAGGAGCTCAGTAATGTCC TCTCAGcc-3′
	Exon 4 cleavage (Bsp*1407I*)	5′-TGAGCACACTGCTGTACAACACCCAC CCATACCcc-3′

Note. Underlined uppercase indicates cleavage site; underlined lowercase indicates unpaired 3′ tail.

PCR and DNA sequencing

Endonuclease cleaved or uncleaved cDNA was used in the PCR reactions. Thermal cycling was conducted with cDNA from 0.2 μg RNA in a final volume of 25 μl containing 10 mM Tris–HCl (pH 8.8), 50 mM KCl, 0.08% NP-40, 1.5 mM $MgCl_2$, and 0.1 mM dNTP each on a Biometro thermal cycler with initial denaturation at 94 °C for 2 min, followed by 30 cycles of 94 °C for 45 s, 62 °C for 45 s, and 72 °C for 60 s for *Fas* gene and of 94 °C for 45 s, 65 °C for 45 s, and 72 °C for 60 s for *PIG-T* gene. PCR products were analyzed on 1% agarose gel. The bands of interest were excised from gels and cloned into pGEM-T vector. Following transformation into *Escherichia coli,* the inserts were sequenced using T7 primer.

Results

RT–PCR serves as an important method for identifying, verifying, and cloning splicing variants. For a specific gene of interest, the level of mRNA varies and RT–PCR may detect only the major variants of high abundance and miss the minor ones of low-level abundance. To illustrate this situation, we mixed two plasmids of different sizes as templates at different ratios and amplified them simultaneously using the same pair of primers. The larger plasmid (pGEM-T) contained a 1026-bp insert of the hnRNP A2 open reading frame (ORF), and the smaller one was the same as the larger one except that the inserted hnRNA A2 had a 264-bp deletion. As is shown in Fig. 2, the smaller template was still detectable when its abundance was 10 times lower than the larger one. However, when its abundance was 100 times lower, it became undetectable and only the product of the larger template was visualized. Cleavage of exon 6 present only in the hnRNP A2 insert in the larger plasmid with *Xho*I endonuclease greatly reduced its PCR product and increased the amplification of the smaller template that produced a detectable band even when it was 100 or 1000 times lower

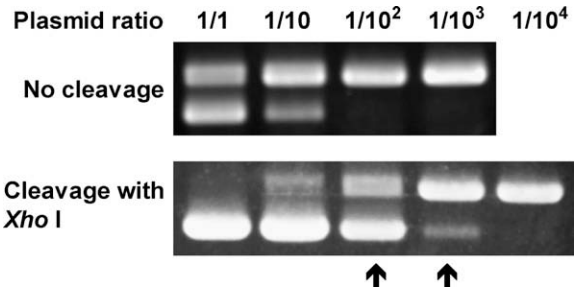

Fig. 2. Simultaneous PCR amplification of two plasmids containing the hnRNP A2 ORF with (smaller) or without (larger) a 264-bp deletion. Here 100 ng larger plasmid (4041 bp) was mixed with 100, 10, 1, and 0.1 ng smaller plasmid (3777 bp), respectively, amplified with or without prior *Xho*I endonuclease digestion (2 h) that cut at exon 6 of the hnRNP A2. Exon 6 was present in the larger plasmid but not in the smaller plasmid. Arrows indicate bands that became detectable after cleavage of exon 6.

than the larger template (Fig. 2, arrows). This result suggests that the conventional PCR can detect splicing variants with changes in abundance within one order of magnitude and that the exon exclusion by endonuclease hydrolysis can increase the detection range by roughly 100-fold.

To test the ability of this method to detect low-abundance mRNA splicing variants, we analyzed the transcripts of *Fas* and *PIG-T* genes. *Fas* gene encodes a cell surface receptor that regulates apoptosis and its signaling pathway. The full-length transcript consists of 9 exons covering 2689 nt. We amplified its transcripts in both MRC5 and HeLa cells, and the results given in Fig. 3 show that the variants are expressed differentially in the two cell types. For the MRC5 cells, four bands were obtained. For the HeLa cells, one major band and one minor band were detected. The largest band in both samples is the full-length transcript of the *Fas* gene. Cleavage of exon 4 of the transcripts cDNA from HeLa cells with *Msp*I prior to PCR led to the detection of two additional bands (Fig. 3, nos. 8 and 10), which were identified, by sequencing, as missing exon 4 and exons 3, 4, and 6, respectively (Table 2).

PIG-T is a subunit of the glycosylphosphatidylinositol (GPI) transamidase multimeric complex localized in the endoplasmic reticulum, and it mediates posttranslational transfer of preformed GPI to proteins bearing a carboxyl-terminal GPI attachment signal [9]. According to the Alternative Splicing Annotation Project (ASAP) database (http://bioinfo.mbi.ucla.edu/ASAP, UniGene Cluster Identifier: Hs.84038), the full-length transcript consists of 12 exons

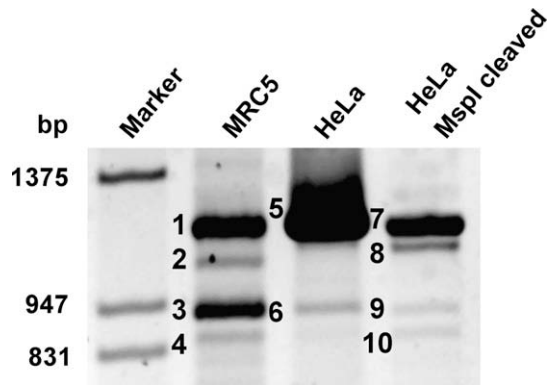

Fig. 3. Identification of *Fas* gene transcripts that are expressed differentially in MRC5 and HeLa cells via exon exclusive RT–PCR. Cleavage of exon 4 of transcripts cDNA from HeLa cells with *Msp*I cleavage (4 h) prior to PCR identified two more bands (8 and 10) that would barely be detected by the conventional PCR. The bands numbered at the left side were cloned into pGEM-T vector and subsequently sequenced. Details of the identified variants are provided in Table 1.

Table 2
Splicing variants of *Fas* gene identified using the exon exclusion method

Number	Amplicon size measured (bp)	Amplicon size actual (bp)	Variant
1	1168	1154	Full-length
2	1030	1016	Δ3
3	913	908	Δ3Δ4
4	840	845	Δ3Δ4Δ6
5	1168	1154	Full-length
6	913	908	Δ3Δ4
7	1168	1154	Full-length
8	1065	1043	Δ4
9	913	908	Δ3Δ4
10	840	845	Δ3Δ4Δ6

Note. cDNA of *Fas* gene was annealed with DNA complementary to exon 4 and subject to cleavage with *Msp*I before PCR amplification.

spanning for 2235 nt, from which five alternative splicing transcripts were predicted (Fig. 4A). When the cDNAs were amplified directly by PCR, only one major band is obvious (Fig. 4B). Endonuclease cleavage of exon 3 with *Sac*I followed by PCR identified at least nine more bands that were barely seen in the direct PCR. When exon 9 was cleaved with *Bsp*407I before PCR amplification, 6 bands were detected in addition to the full-length transcript. We sequenced 11 bands, and the result is shown in Table 3. Four new splicing variants (bands 4, 6, 7, and 11) were found in addition to the five reported ones.

Discussion

Amplification of mRNA sequences by RT–PCR plays an important role in gene discovery and gene expression analysis [6]. Successful detection of an alternative splicing variant depends on its relative abundance among the variants that are amplified simultaneously. As is demonstrated in Fig. 2, a template might not be detectable when its level is 100 times lower than the other template due to the competition for polymerase, primers, and dNTP. By cleaving the high-abundance template, the detection of the low-abundance template lacking the cleavage site by PCR is enhanced by approximately 100-fold, making it detectable even when its level is 1000 times less abundant than its competing partner. This protocol should provide a useful tool for amplifying and identifying low-abundance alternative splicing variants. For a gene of interest, cleaving a specific exon in the cDNA from the dominant high-abundance splicing variants prevents them from being amplified by PCR and, thus, greatly facilitates the amplification of the low-abundance variants lacking the cleaved exons.

The analysis of the two example genes demonstrated the effectiveness of our protocol in identifying both known and unknown splicing variants. The transcripts that can be detected will still depend on their relative abundance in the cleaved samples, as in the untreated samples. For this reason, choosing different exon exclusion may help to detect more variants. For instance, cleavage for the *PIG-T* gene with *Bsp*1407I identified two variants (nos. 9 and 11*) that were not detected in the cleavage with *Sac*I (Fig. 4). In the former case, the cleavage of exon 9 detected two isoforms (Table 3, nos. 5 and 11*) in which exon 3 was skipped. However, cleavage of exon 3 failed to detect isoform 11*; this might be caused by the presence of other variants at higher abundance (e.g., nos. 3 and 6*) or the nearby band (no. 6*) in the gel. Therefore, a combination of different exon cleavages should result in effective improvement of detection spectrum. Cleaving more than one exon or using more than one endonuclease may also improve detection sensitivity.

Because the cleaved exons must be placed between the two primers for PCR, our method is not suitable for identifying splice variants lacking the first or last exon. Restriction endonucleases

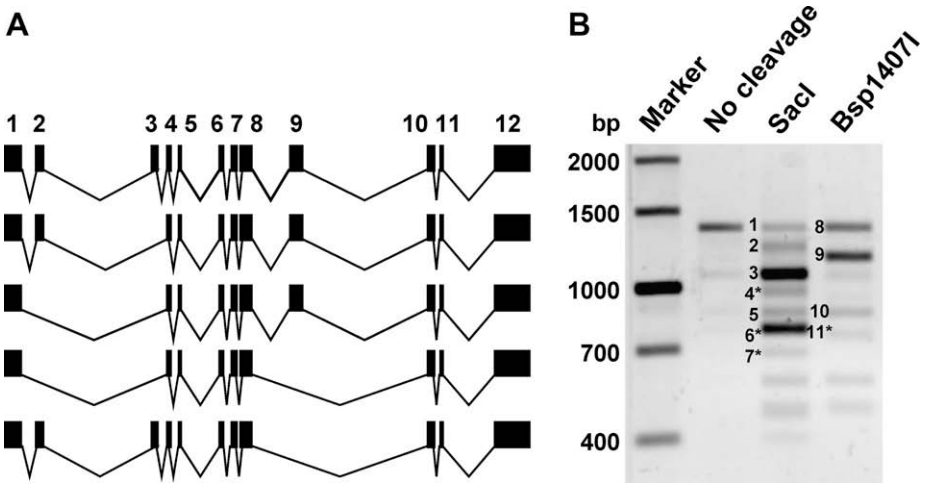

Fig. 4. (A) Reported alternative splicing transcripts of *PIG-T* gene in the ASAP database. (B) Identification of *PIG-T* gene transcripts from HeLa cells via exon exclusive RT–PCR. cDNAs were cleaved at exon 3 with *Sac*I (4 h) or at exon 9 with *Bsp*1407I (10 h) before PCR amplification. Exclusion of exons 3 and 9 visualized additional bands that were undetectable by the conventional PCR. The bands numbered at the left side were cloned into pGEM-T vector and subsequently sequenced. Details of the identified variants are provided in Table 2.

Table 3
Splicing variants of *PIG-T* gene identified using the exon exclusion method

Number	Exon cleaved	Amplicon size measured (bp)	Amplicon size actual (bp)	Variant
1		1370	1365	Full-length
2	3	1230	1237	Δ3
3	3	1025	1059	Δ2Δ3
4*	3	934	958	Δ2Δ3Δ4
5	3	821	858	Δ2Δ3Δ9
6*	3	743	783	Δ2Δ3Δ4Δ5Δ6
7*	3	663	679	Δ2Δ3Δ4Δ5Δ6Δ7
8		1376	1365	Full-length
9	9	1158	1164	Δ9
10	9	821	858	Δ2Δ3Δ9
11*	9	742	753	Δ2Δ3Δ4Δ9

* Newly identified isoform that is not reported in the ASAP database. cDNA of *PIG-T* gene was annealed with DNA complementary to exon 3 or 9 and subject to cleavage with *Sac*I or *Bsp*1407I, respectively, before PCR amplification.

were used in this study, which relies on the presence of restricting sites on the targeted exons. The availability of a duplex-specific nuclease [10,11] can remove this limitation and allow digestion of any exons. Recently, a method for enrichment of mRNA isoform with specific alternative splicing, using negative selection to suppress amplification of unwanted variants or positive selection for desired variants, was reported. Our method is similar to the negative selection in that report where unwanted transcript was destroyed by RNase H treatment and desired transcript was enriched by affinity separation [12]. The enhancement of detection of specified variant in that work was 16- to 19-fold using single selection and 26-fold when both negative and positive selections were combined. Compared with that method, our method is much more sensitive and simple. Although our method lacks the ability to selectively enrich variants carrying a specific exon, this will not limit its potential when the full-length information of the target mRNA is available and duplex-specific nuclease can be used to cleave each exon. We expect that our method will provide an effective and flexible tool for identifying and profiling alternative splicing variants when one is interested in a specific gene.

Acknowledgments

This work was supported by grant 2007CB507402 from MSTC and by grants 20572082 and 30670451 and the Science Fund for Creative Research Groups from NSFC.

References

[1] B. Modrek, C. Lee, A genomic view of alternative splicing, Nat. Genet. 30 (2002) 13–19.
[2] J.M. Johnson, J. Castle, P. Garrett-Engele, Z. Kan, P.M. Loerch, C.D. Armour, R. Santos, E.E. Schadt, R. Stoughton, D.D. Shoemaker, Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays, Science 302 (2003) 2141–2144.
[3] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, et al., Initial sequencing and analysis of the human genome, Nature 409 (2001) 860–921.
[4] J.O. Bishop, J.G. Morton, M. Rosbash, M. Richardson, Three abundance classes in HeLa cell messenger RNA, Nature 250 (1974) 199–204.
[5] M. Blaustein, F. Pelisch, A. Srebrow, Signals, pathways, and splicing regulation, Intl. J. Biochem. Cell Biol. 39 (2007) 2031–2048.
[6] J.W. Larrick, Message amplification phenotyping (MAPPing): principles, practice, and potential, Trends Biotechnol. 10 (1992) 146–152.
[7] P. Chomczynski, N. Sacchi, The single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction: Twenty-something years on, Nat. Protoc. 1 (2006) 581–585.
[8] P. Bauer, A. Rolfs, V. Regitz-Zagrosek, A. Hildebrandt, E. Fleck, Use of manganese in RT–PCR eliminates PCR artifacts resulting from DNase I digestion, BioTechniques 22 (1997) 1128–1132.
[9] K. Ohishi, K. Nagamune, Y. Maeda, T. Kinoshita, Two subunits of glycosylphosphatidylinositol transamidase, GPI8 and PIG-T, form a functionally important intermolecular disulfide bridge, J. Biol. Chem. 278 (2003) 13959–13967.
[10] V.E. Anisimova, D.V. Rebrikov, P.A. Zhulidov, D.B. Staroverov, S.A. Lukyanov, A.S. Shcheglov, Renaturation, activation, and practical use of recombinant duplex-specific nuclease from Kamchatka crab, Biochemistry 71 (2006) 513–519.
[11] Y. Zhao, H. Hoshiyama, J.W. Shay, W.E. Wright, Quantitative telomeric overhang determination using a double-strand specific nuclease, Nucleic Acids Res. 36 (2008) e14.
[12] N.Y. Wang, H.C. Lai, L.C. Au, Methods for enrichment of a mRNA isoform with specific alternative splicing, Anal. Biochem. 359 (2006) 189–193.